

GenCore version 5.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 16:43:39 ; Search time 2168 Seconds
(without alignments)
9453.747 Million cell updates/sec

Title: US-09-484-577A-3

Perfect score: 501

Sequence: 1 actctccagcctctcacgca.....cagaatgcgcatgatcat 501

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rod:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	135.8	27.1	302650	1	AP005958	AP005958 Bradyrhiz
C 2	97.2	19.4	10902	1	AE011858	AE011858 Xanthomon
C 3	92.4	18.4	3526	1	AF273216	AF273216 Rhizobium
C 4	88.8	17.7	28804	1	SSU51197	SSU51197 Sphingomona
C 5	88.8	17.7	28804	6	AR068625	AR068625 Sequence
C 6	87.6	17.5	302022	1	AE012558	AE012558 Xylella f
C 7	86.8	17.3	15552	1	AE004049	AE004049 Xylella f
C 8	86	17.2	1428	1	NME331263	AJ391263 Neisseria
C 9	86	17.2	4254	1	AF121772	AF121772 Neisseria
C 10	86	17.2	7824	1	NME331260	AJ391260 Neisseria
C 11	86	17.2	11381	1	AE002524	AE002524 Neisseria
C 12	86	17.2	326301	1	NMA622491	AL162757 Neisseria
C 13	86	17.2	349980	6	AX044033	AX044033 Sequence
C 14	86	17.2	349980	6	AX044034	AX044034 Sequence
C 15	80.6	16.1	301708	1	AE016792	AE016792 Pseudomon
C 16	80.4	16.0	189050	1	AL646077	AL646077 Ralstonia
C 17	69.8	13.9	11180	1	AE013811	AE013811 Yersinia
C 18	69.8	13.9	334050	1	AJ414151	AJ414151 Yersinia
C 19	67.4	13.5	298600	1	AP005959	AP005959 Bradyrhiz
C 20	62.6	12.5	6645	1	BPCYADE	X14199 B.pertussis
C 21	56.8	11.3	7493	1	AF193064	AF193064 Caulobact
C 22	56.8	11.3	13759	1	AF005779	AF005779 Caulobact
C 23	56.8	11.3	19883	1	AF062345	AF062345 Caulobact
C 24	54.2	10.8	303550	1	SCO939118	AL939118 Streptomy
C 25	53.2	10.6	208050	1	AL646083	AL646083 Ralstonia
C 26	52.6	10.5	4050	1	AV158231	AV158231 Caulobact
C 27	52	10.4	2000	6	AX655393	AX655393 Sequence
C 28	51.6	10.3	5120	1	APCLYI	X61112 A.pleuropne
C 29	51.6	10.3	8292	1	APHLXIAD	X68595 A.pleuropne
C 30	51.6	10.3	8370	6	AR091872	AR091872 Sequence
C 31	51.6	10.3	8370	6	AR268551	AR268551 Sequence
C 32	51.2	10.2	300425	1	AP005038	AP005038 Streptomy
C 33	48.6	9.7	348411	1	AP003007	AP003007 Mesorhizo
C 34	48.4	9.7	4247	1	ACNAPPAB	M65808 Actinobacil
C 35	48	9.6	7413	1	PAAPRAPG	X64558 P.aeruginos
C 36	48	9.6	11588	1	AE004554	AE004554 Pseudomon
C 37	47.8	9.5	6400	6	E30060	E30060 ABC Transpo
C 38	47.8	9.5	11273	1	AB023289	AB023289 Pseudomon
C 39	47.6	9.5	11383	1	AE004761	AE004761 Pseudomon
C 40	47.2	9.4	299350	1	SME591786	AE006947 Mycobacte
C 41	47	9.4	18857	1	AE006947	AE006947 Mycobacte
C 42	47	9.4	35420	1	MTCY22G10	Z84724 Mycobacteri
C 43	47	9.4	42741	1	MSGY423	AD000014 Mycobacte
C 44	47	9.4	324050	1	BX248335	BX248335 Mycobacte
C 45	46.8	9.3	76093	1	AF416330	AF416330 Ruegeria

ALIGNMENTS

RESULT 1
AP005958/c
LOCUS
DEFINITION Bradyrhizobium japonicum USDA 110 DNA linear BCT 28-MAR-2003
24/31.
ACCESSION AP005958 BA000040
VERSION AP005958.1 GI:27354550
KEYWORDS
SOURCE Bradyrhizobium japonicum USDA 110
ORGANISM Bradyrhizobium japonicum USDA 110
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE
AUTHORS Kaneke, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T.,
Saeamoto, S., Watanabe, A., Idesawa, K., Iriguchi, M., Kawashima, K.,

Kohara, M., Matsumoto, M., Shimpō, S., Tsuruoka, H., Wada, T., Yamada, M. and Tabata, S.
Complete genomic sequence of nitrogen-fixing symbiotic bacterium Bradyrhizobium japonicum USDA110
DNA Res. 9 (6), 189-197 (2002)
22484998
MEDLINE
PUBMED
12597275
REFERENCE
2
AUTHORS
Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Igesawa, K., Iriguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimpō, S., Tsuruoka, H., Wada, T., Yamada, M. and Tabata, S.
Complete genomic sequence of nitrogen-fixing symbiotic bacterium Bradyrhizobium japonicum USDA110 (supplement)
DNA Res. 9 (6), 225-256 (2002)
22485002
MEDLINE
PUBMED
12597279
REFERENCE
3 (bases 1 to 302650)
AUTHORS
Kaneko, T.
TITLE
Direct Submission
JOURNAL
Submitted (20-NOV-2002) Takakazu Kaneko, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research; 2-6-7 Kazusa-kamatori, Kisarazu, Chiba 292-0812, Japan
(E-mail: kaneko@kazusa.or.jp,
URL: http://www.kazusa.or.jp/rhizobase/,
Tel: 81-438-52-3935 (ex. 2338), Fax: 81-438-52-3934)
Location/Qualifiers
1..302650
/organism="Bradyrhizobium japonicum USDA 110"
/mol_type="Genomic DNA"
/strain="USDA110"
/db_xref="taxon:224911"
31..1110
/gene="blr6272"
31..1110
/gene="blr6272"
/note="AraC family
ORF ID:blr6272"
/codon_start=1
/transl_table=11
/product="transcriptional regulatory protein"
/protein_id="BAC51537.1"
/db_xref="GI:27354551"
/translation="MENLLESVLESFGLGLPVAAGORDAKRGAMFTDLOAPAWID LOAEPDHOLOAOATPRERWRIDRGPEAAIEDITISRWEDFRDSSSHDATTPEDE RYFVGIALKATRAKLTEDROVIEDGTPAGTLVUSAPSKPLAAQFAPCAFLHFLISA DHPAQSPAEGDLVLLKDPDLAELARALTEHGDAOREFTRCVGQILAMHLARLE PPARVVALPFWLRVVEQITIAUNDFRCISLSELANVAGLSRMHFAAQFTATGTGYRPR EYLNHRIEHAKTLATTTGRPLAELIALAVGFSTQAHFSTVFKRISGQSPARWELASKA EPLAAEALPRRPAADRDMMASAAA"
complement (1428..3794)
/gene="blr6273"
complement (1428..3794)
/gene="blr6273"
/note="ORF ID:blr6273
probable biotin sulfoxidase"
/codon_start=1
/transl_table=11
/protein_id="BAC51538.1"
/db_xref="GI:27354552"
/translation="MDRYATATMYRNKQDMDDTIGFPDPLDLSGDKPKHTSHGWFS ARNGAGLEVRAYAGDPDNGIINDFPGLRHQARIQAPAIRRGWLERGEGPDRRGR DEFVSWEKALDILGDELGRVIDTRGPGAVFGSGYMSAGRPFAHQSOVHRFLNIA MGVYRVSNSYSSGASVLLPQILAGVEDITKRNVTWEOJATETDILVAFPGMALONS MVAGGSISKVRESBAARKGCEFLVSPLEDLPVEAGAEWTCVSTDTMLMG IIVTLVSEKHQDAFLDRYTEGPVFVLYLTISGDSGQSAEAAISIGVADTITRL ARRLAGKEALITSHSLRABHEGQPVMMGMVLAALQIGLPGGTYAYSLGAIYGY RVNVDVPGTLGQNGVRDIFIVARIADMLNPGATRYNGETRTYPOIRLYVWAGG NPFFHQDINLRKARQOLDTFVYVHELAWTATARHADIVLPSTWMLEREDIGYSSNDP LVMAMHRIABEPGLARDYEIFADLARIAGLRARPEGTGRSRLKHLVEPTASLAM RGLAEPSPDEFWQBSLVVQQPDGGRRLRFREDPVAAHALPTPSGRILFSTRIAGH GDADCPGHWLDKDKMPKAGPCFLVANAQPVTLRSLQSLDFGGHSLGAKHRGHEVARM NPVDARGINDGSDIILFNERGACLAHVHTDGIAPGVVQLFTGAWYDMPDDEDDAP

LCVGNPNVLTRDYGTSSFAOGCTQLTVEVEKFTGNLPPIRAPDPV"
complement (3898..4947)
/gene="blr6274"
complement (3898..4947)
/gene="blr6274"
/note="ORF ID:blr6274
putative dehydrogenase"
/codon_start=1
/transl_table=11
/protein_id="BAC51539.1"
/db_xref="GI:27354553"
/translation="MTGASDTVAIGADALRPFPAADLLRAGGFTADQADETAAMLVWAN LRGADSHGLRIPRYVEMVELGLINPAEPQVSGKGAVALIDQAPGAVGMSLAAR KASELADTHGICMCVARSISAGATGYVAHQVQAQRIGIVMTASKPLMIYHSSAEG VSTNPJIAAFTGPNDRPLLDMSTASVALNKIQAARDAGQPIPVGHAVDNGQATTD PREKAVILPMAGPKPGSGLSLMIEVLASVLTIAPLIALVALVHGDPGGNGLVAVLDPSA FGDNFTLEDRLGAAIKSLPPAAGADSVYLPFGRGFEEMERRSRAGIPLARGTIVSRLL ALASKLNAVPEGVA"
complement (4950..5873)
/gene="blr6275"
complement (4950..5873)
/note="ORF ID:blr6275
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC51540.1"
/db_xref="GI:27354554"
/translation="MSAPFLSQPTREIPMQBKHTFRSDGLKLSAVLHVDFDARKPQRL PAFIVLHGFGVSKDSEHAEIOARMLQMGYVALRDFRCCGESGERAQVRCFDQVAD AKNALTFLAERREVDPGRIGVGVHSGFGAAVVAAGVDDRFACVYIKSSCGWDGKERKFR QGHPTPEAWAKFTSLLEKRGKQKQETGOSMVISRFDVAPIPEHLRKNLSPKAIIMEIPV ETAMSMYFRADEVVGNTPAPLILFHPANDTITPTEQSIIRKFEKAGQPAELVLITGT SHFPLAEGDAPRTKALIKGHLDKFPFSPPLGA"
complement (5910..6812)
/gene="blr6276"
complement (5910..6812)
/gene="blr6276"
/note="ORF ID:blr6276
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAC51541.1"
/db_xref="GI:27354555"
/translation="MPPIVHAVGHVKINTTAIEAVVRDATEILGLQITYTDEGQTWLS SNGRAELVLRSHENATHIGLEALTVDAREASRVESAGCRVISRPSLDCMNAAG VTFTDPEGLREVEHTPTDQIYGRYPYTHGVGNRMDHLNLTSPDPIATROOLERICG LRIISRMVDDLSLMYGGNRQHIVGVYRGVGLHLYSFEFKBFNDYLLKGLDLDRFD RQMLWGPGRHFGONTYATYTTDSSGAMIECSGSMALIADDDGFPVNVITNLERPGNVR CMNVWGTAPLEWRQHFHPTAIT"
7080..7853
/gene="blr6277"
7080..7853
/gene="blr6277"
/note="GntR family
ORF ID:blr6277"
/codon_start=1
/transl_table=11
/product="transcriptional regulatory protein"
/protein_id="BAC51542.1"
/db_xref="GI:27354556"
/translation="MILTGSRMAGRRMRERVKPAGETRAADVLRHMRADIISCTLK PGAKRLPALRDIIVASFSPTLREALSLVAGSLVTAEDQRCFVAVPVISIDLNDLITYY RVLTRECVALKNGDADWEADIIIGA PHMDRLQNLGNSYLYSEEWKLGHPHFS LVAACGSNLEIKOKLFEHARHTRMSSQFRTKWRADKVEHKMIMDSVVARDAKAE ELIERHIFETDQNVIKHAGHLFVTSDEERPRTRSRVAAE"
complement (7993..8880)
/gene="ilvG"
complement (7993..8880)
/gene="ilvE"
/note="ORF ID:blr6278"
/codon_start=1
/transl_table=11

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OV nucleic - nucleic search, using sw model

Run on: October 23, 2003, 17:23:09 ; Search time 1656 Seconds
(without alignments)
7352.985 Million cell updates/sec

Title: US-09-484-577A-3
Perfect score: 501
Sequence: 1 actctccagctctcacgga.....cagaatgcgggatgatcat 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hct:*
9: gb_est1:*
10: gb_est2:*
11: gb_hct:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	60.6	12.1	932	29	CNS0070E AL066254 Drosophil
2	55	11.0	925	29	CNS0091P AL053013 Drosophil
3	49.4	9.9	812	29	BZ549462 pacs1-60-
C 4	45.2	9.0	1148	13	BX374895 BX374895

C 5	44.8	8.9	508	9	AV630811
C 6	43.6	8.7	772	28	AF075872
C 7	42	8.4	839	29	CNS004NE
C 8	41.4	8.3	612	14	CB214812
C 9	41.4	8.3	620	14	CB211981
C 10	41	8.2	1201	13	BX360624
C 11	40.8	8.1	932	29	CNS0072Q
C 12	40.8	8.1	1201	13	BX366664
C 13	40.6	8.1	420	9	AJ474950
C 14	40.6	8.1	426	13	BX437276
C 15	40.6	8.1	835	29	BZ557064
C 16	40.6	8.1	851	29	BZ551811
C 17	40.6	8.1	919	13	BQ890839
C 18	40.6	8.1	927	29	BZ579186
C 19	40.4	8.0	774	12	B1956591
C 20	40.2	8.0	478	13	B0980310
C 21	40.2	8.0	504	12	BJ450136
C 22	40.2	8.0	531	14	CD054442
C 23	40.2	8.0	581	9	AV921113
C 24	40.2	8.0	586	9	AV913605
C 25	40.2	8.0	588	12	BJ468187
C 26	40.2	8.0	605	9	AV918908
C 27	40.2	8.0	626	12	BJ473505
C 28	40.2	8.0	628	12	BJ457563
C 29	40.2	8.0	638	14	CB880020
C 30	40.2	8.0	659	9	AV931532
C 31	40.2	8.0	660	9	AV946608
C 32	40.2	8.0	660	12	BJ465650
C 33	40.2	8.0	660	12	BJ475855
C 34	39.8	7.9	588	13	B036564
C 35	39.8	7.9	603	13	BQ279904
C 36	39.8	7.9	627	13	BU092391
C 37	39.8	7.9	633	13	BU098243
C 38	39.8	7.9	639	9	A1770858
C 39	39.6	7.9	574	12	BM349363
C 40	39.6	7.9	576	9	A1881920
C 41	39.6	7.9	577	12	BM267007
C 42	39.6	7.9	643	13	BU499802
C 43	39.6	7.9	645	29	CNS01213
C 44	39.6	7.9	673	29	BZ563812
C 45	39.6	7.9	686	12	B1959511

ALIGNMENTS

RESULT 1
CNS0070E/c

LOCUS

DEFINITION

AL066254

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CNS0070E 932 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR14D21 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

AL066254

AL066254.1 GI:4945121

GSS.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

Genoscope.

1 (bases 1 to 932)

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoss in Pieter de Jong's laboratory in the Department of

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

DNM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 16:42:52 : Search time 232 Seconds
(without alignments)
5829.390 Million cell updates/sec

Title: US-09-484-577A-3

Perfect score: 501

Sequence: 1 actctccagctctcacgca.....cagaatggcgatgatcat 501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_13Jun03.*

1:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1980.DAT.*
2:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1981.DAT.*
3:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1982.DAT.*
4:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1983.DAT.*
5:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1984.DAT.*
6:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1985.DAT.*
7:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1986.DAT.*
8:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1987.DAT.*
9:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1988.DAT.*
10:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1989.DAT.*
11:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1990.DAT.*
12:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1991.DAT.*
13:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1992.DAT.*
14:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1993.DAT.*
15:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1994.DAT.*
16:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1995.DAT.*
17:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1996.DAT.*
18:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1997.DAT.*
19:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1998.DAT.*
20:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA1999.DAT.*
21:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA2000.DAT.*
22:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA2001A.DAT.*
23:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA2001B.DAT.*
24:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA2002.DAT.*
25:	/SIDSI/gcgdata/geneSeq/geneSeq-n-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	88.8	17.7	28804	17 AAT37329	Sphingan biosynthe
C 2	88.8	17.7	28804	18 AAT92474	Sphingomonas genus
C 3	88.8	17.7	28804	20 AAV99812	Sphingomonas S88 s
C 4	88.8	17.7	28804	20 AAV81474	Chromosomal fragme
5	86	17.2	1305	21 AAZ54336	Neisseria meningit
6	86	17.2	1428	21 AAZ54335	Neisseria meningit
7	86	17.2	24158	21 AAAB1532	N. meningitidis pa
C 8	86	17.2	349980	21 AAF21611	Neisseria meningit

C 9	86	17.2	349980	21 AAF21612	Neisseria meningit
C 10	86	17.2	837096	21 AAAB1489	N. meningitidis pa
11	51.6	10.3	5120	18 AAT73218	Ap1BD gene. Acti
12	51.6	10.3	5120	21 AAZ88585	A. pleuropneumonia
13	47.8	9.5	6400	20 AAZ22701	Pseudomonas fluore
14	47	9.4	4403765	22 AAI99683	Mycobacterium tube
15	47	9.4	4411529	22 AAI99682	Mycobacterium tube
16	45	9.0	7184	15 AAQ70050	ptx gene of Pasteu
17	44.8	8.9	349980	24 ABQ81842	Bifidobacterium lo
18	43.4	8.7	7183	11 AAQ06074	Sequence encoding
19	43	8.6	24379	18 AAT93095	Streptomyces freno
20	43	8.6	24379	19 AAQ25925	Streptomyces roseo
C 21	42	8.4	4403765	22 AAI99683	Mycobacterium tube
C 22	42	8.4	4411529	22 AAI99682	Mycobacterium tube
C 23	40	8.0	45055	25 ABZ66808	Orthomyxovirin biosy
24	39.8	7.9	1320	25 ABX56068	M. echinospira cal
C 25	39	7.8	1761	22 AAF66404	C glutamicum codin
C 26	39	7.8	1998	22 AAF72013	Corynebacterium gl
27	39	7.8	349980	22 AAH68528	C glutamicum codin
28	38.2	7.6	5802	24 ABS78696	S. kaniharaensis D
29	37.6	7.5	1553	23 AAS89358	DNA encoding novel
C 30	37.6	7.5	2931	23 AAS92554	DNA encoding novel
C 31	37.6	7.5	10732	21 AAAL0594	Gene encoding a su
C 32	37.2	7.4	375	25 ABX48203	Bovine EST associa
33	37.2	7.4	1194	17 AAT58555	Streptomyces prist
34	37.2	7.4	4496	17 AAT58553	Streptomyces prist
C 35	37.2	7.4	5857	21 AAAS8471	Nucleotide sequenc
36	37.2	7.4	76804	24 ABS78942	E. coli CFT073 gen
37	37.2	7.4	349980	24 ABQ81847	Bifidobacterium lo
C 38	37	7.4	1951	19 AAV11459	C. acidivorans gam
C 39	37	7.4	1981	21 AAZ50482	Corn sulphate perm
40	36.6	7.3	276	24 ABN75424	Human transport pr
41	36.4	7.3	1408	24 ABK53812	Rat sequence diffe
42	36.4	7.3	1739	22 ADI9384	Rat CCAAT/enhancer
43	36.4	7.3	1739	23 ABK51406	DNA encoding rat C
44	36.4	7.3	1739	24 ABK44082	Rat C/EBP-beta pro
45	36.4	7.3	1739	24 ABK51998	DNA encoding wild-

ALIGNMENTS

RESULT 1
AAT37329/c
ID AAT37329 standard; DNA; 28804 BP.
XX AAT37329;
AC AAT37329;
XX 30-NOV-1996 (first entry)
DT Sphingan biosynthetic gene region.
DE Sphingan; polysaccharide; speB gene; glucosyl-IP-transferase; ds.
XX Sphingomonas strain S88 (ATCC 31554).
OS
XX
XX Key Location/Qualifiers
FT CDS complement (1942..1944)
FT /*tag= a
FT /codon_start= 1942..1944
FT /note= "speB gene putative initiation codon"
FT complement (3311..3313)
FT /*tag= b
FT /codon_start= 3311..3313
FT /note= "speB gene putative initiation codon"
FT complement (5323..5325)
FT /*tag= c
FT /codon_start= 5323..5325
FT /note= "speB gene putative initiation codon"
FT 5526..5528
FT /*tag= d
FT /codon_start= 5526..5528
FT /note= "speB gene putative initiation codon"

```

CDS      complement (7076..7078)
         /tag= e
         /codon_start= 7076..7078
         /note= "spSI gene putative initiation codon"
         7588..7590
         /tag= f
         /codon_start= 7588..7590
         /note= "spSK gene putative initiation codon"
         8643..8645
         /tag= g
         /codon_start= 8643..8645
         /note= "spSL gene putative initiation codon"
         complement (10938..10940)
         /tag= h
         /codon_start= 10938..10940
         /note= "spSJ gene putative initiation codon"
         11569..11571
         /tag= i
         /codon_start= 11569..11571
         /note= "spSF gene putative initiation codon"
         12886..12888
         /tag= j
         /codon_start= 12886..12888
         /note= "spSD gene putative initiation codon"
         15165..15167
         /tag= k
         /codon_start= 15165..15167
         /note= "spSE gene putative initiation codon"
         15883..15885
         /tag= l
         /codon_start= 15883..15885
         /note= "urf32 gene putative initiation codon"
         16748..16750
         /tag= m
         /codon_start= 16748..16750
         /note= "urf28 gene putative initiation codon"
         complement (21082..21084)
         /tag= n
         /codon_start= 21082..21084
         /note= "atrB gene putative initiation codon"
         21706..23118
         /tag= o
         /label= spSB
         /product= glucosyl IP-transferase
         23238..23240
         /tag= p
         /codon_start= 23238..23240
         /note= "rnsA gene putative initiation codon"
         24113..24115
         /tag= q
         /codon_start= 24113..24115
         /note= "rnsC gene putative initiation codon"
         24683..24685
         /tag= r
         /codon_start= 24683..24685
         /note= "rnsB gene putative initiation codon"
         25744..25746
         /tag= s
         /codon_start= 25744..25746
         /note= "rnsD gene putative initiation codon"
         complement (27534..27536)
         /tag= t
         /codon_start= 27534..27536
         /note= "urf31 gene putative initiation codon"
         27747
         /tag= u
         /codon_start= 27747..27749
         /note= "urf34 gene putative initiation codon"

EP728841-A2..
28-AUG-1996..

```

```

PF 24-JAN-1996; 96EP-0300467.
XX
XX
PR 24-JAN-1995; 95US-0377440.
XX
XX
PA (SHIN-) SHINETSU BIO INC.
PA (SHIE ) SHINETSU CHEM CO LTD.
XX
XX Armentrout RW, Mikolajczak M, Pollock TJ, Thorne L;
XX Yamazaki M;
XX
XX WPI; 1996-386292/39.
DR P-PSDB; AAW03997.
XX
XX
XX New isolated DNA from Sphingomonas sp. - used for transforming
PT recipient bacteria to obtain hyper-producers of sphing
PT polysaccharide(s).
XX
XX
PS Claim 32; Page 56-70; 105pp; English.
XX
XX A 28.8 kb chromosomal fragment of Sphingomonas strain S88 was
CC isolated on the basis of its ability to restore sphing
CC biosynthetic capability to Sphingomonas mutant S88m260. It
CC contains 23-25 genes, including spa genes coding for biosynthesis of
CC the polysaccharide sphing, rhs genes coding for dTDP-(L)ribose
CC biosynthesis, atrB genes coding for a transport function and some
CC unidentified open translation reading frames (urf). The spaB gene
CC was identified that is believed to code for glucosyl IP-transferase
CC (AAW03997), an enzyme catalysing the first step of assembly of
CC sphing carbohydrites. DNA fragments of S88 can be inserted into
CC a vector in multiple copies and used to produce engineered bacteria
CC that are hyper-producers of sphing.
XX
XX Sequence 28804 BP; 4974 A; 9806 C; 9228 G; 4796 T; 0 other;
XX
XX
XX Query Match 17.7%; Score 88.8; DB 17; Length 28804;
XX Best Local Similarity 56.5%; Pred.No.9.4e-13;
XX Matches 165; Conservative 0; Mismatches 127; Indels 0; Gaps 0;
XX
QY 80 TTCCTCGCGCGCCCTCGAAATTGCGAGACGCCCATCTCCACCGCGAGACTCAGC 139
DB 18808 TTCTCTCGCGCGCGCTCGAGATCATCGAACGTCCGGTGTGCCCGCCGCGCTCACC 18749
QY 140 GCGCGCBTTGCTTGCTGCTGCTTGTACTGCGCGCTGGCGTGGCGGGTCTCGGCAGGATC 199
DB 18748 GCGCGGGTGATGTGCGCGGGCTGGCGATCACCGCTGCTGCAATCGGCGCGTG 18689
QY 200 GACATGTTGCTTCTGCATCCAGAAAGATCGTGGCGGGGACCGTGTAAAGCTGTTGAG 259
DB 18688 GAAGTGGTGGCGCGCAGCAGCGCGCGCATCGCCCGCATCGGCGAGACCAAGATCGTG 18629
QY 260 CCCTCGAGTGGCGGTGGTGGCGGCCTCATGTCCGGATGGCCAAACCGTCAAGGCC 319
DB 18628 TCCCGGAAAGCGGTATCGTCGGCGCATCTTGGTGGCGAGGGCGAGAGTTTCAAG 18569
QY 320 GCGCAGATTCTGATCGAGCTGGATCCATTCGCGGGTGGTGTGGATGTTGCGC 371
DB 18568 GGCGAGTGTGATCAGCTCGATCCACCATGTCCGCGAGCGGAAGCGCGC 18517
XX
XX
XX RESULT 2
XX AAT92474/c
XX ID AAT92474 standard; DNA; 28804 BP.
XX
XX AC AAT92474;
XX
XX
XX 04-FEB-1998 (first entry)
XX
XX
XX Sphingomonas genus microbe isolated DNA sequence producing sphing.
XX
XX Sphingomonas microbe; sphing polysaccharide biosynthesis gene;
XX sphing S-88; spaB gene; ss.
XX
XX Sphingomonas sp.

```


Db 223 GTTGACAGATTGACATTTGTTCTACAGCTTACAGTAAATTTCTTTCAGGTAGCCGTAGC 282
QY 248 AAGCTGGTTACGCCCTCGAGGTGGGTGGTGGCGCCACTGATGTCGCGATGCCRA 307
Db 283 AAGATATTCAATCTTTGGAAACGAGTAGTTAAAGCAGTTTATGTCGTTGATGTCRA 342
QY 308 ACCGTCAAGCGCGGAGAGATCTGATCGAGCTGGATCCATTCGCGGATGTTGGATGTT 367
Db 343 AATGTTCAACAGGTGAATATTTAGTAGATTAGTGGGAATCGGTTTCAGATAGTGATGTT 402
QY 368 GCGCCCGCTCAGAGTCCAT 387
Db 403 GCTCAGTCGGAAGACCC 422

RESULT 2

US-10-156-761-4249/c
; Sequence 4249, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156-761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 4249
; LENGTH: 1803
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1803)
US-10-156-761-4249

Query Match 10.2%; Score 51.2; DB 14; Length 1803;
Best Local Similarity 69.4%; Pred. No. 2.4e-05;
Matches 100; Conservative 0; Mismatches 38; Indels 6; Gaps 2;
QY 333 TCGAGCTGATCCATTCGCGGTGGTGGATGTTGCGCCCGTC-AGAGTCCATCAGC 391
Db 764 TTGCGCAGGACCCATTCGCGGTGGTGGATGTTGCGCGCGGTGGACAGGTCCATGACG 705
QY 392 GTGTGCGCGCCCGGAGATCGCCACCATCTTGTGACCTTTTCTTCCACCGACG----A 446
Db 704 GTGTGCGCGCCCGGAGTCCGCCAGGTCTCTTCGACCTCTCTTCGATGAGGAA 645
QY 447 GTACCGCGGAGTTCGGATATTG 470
Db 644 GTACCGCGGAGTTCGGATATTG 621

RESULT 3

US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156-761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (418715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1
Query Match 10.2%; Score 51.2; DB 14; Length 9025608;
Best Local Similarity 69.4%; Pred. No. 3.9e-05;
Matches 100; Conservative 0; Mismatches 38; Indels 6; Gaps 2;
QY 333 TCGAGCTGATCCATTCGCGGTGGTGGATGTTGCGCCCGTC-AGAGTCCATCAGC 391
Db 5232924 TTGCGCAGGACCCATTCGCGGTGGTGGATGTTGCGCGCGGTGGACAGGTCCATGACG 5232983
QY 392 GTGTGCGCGCCCGGAGATCGCCACCATCTTGTGACCTTTTCTTCCACCGACG----A 446
Db 5232984 GTGTGCGCGCCCGGAGTTCGCCAGGTCTCTTCGACCTCTCTTCGATGAGGAA 5233043
QY 447 GTACCGCGGAGTTCGGATATTG 470
Db 5233044 GTACCGCGGAGTTCGGATATTG 5233067

RESULT 4

US-10-156-761-4175
; Sequence 4175, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156-761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 4175
; LENGTH: 819
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(819)
US-10-156-761-4175

Query Match 8.7%; Score 43.8; DB 14; Length 819;
Best Local Similarity 45.9%; Pred. No. 0.002;
Matches 150; Conservative 0; Mismatches 177; Indels 0; Gaps 0;
QY 62 ACCCGGAGCAGCTGGCCTTCCTGCGCGCGCCTCGAAATTTGTGAGACGCGGCATCT 121

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 17:26:14 ; Search time 58 Seconds
(without alignments)
3812.637 Million cell updates/sec

Title: US-09-484-577A-3

Sequence: 1 actctcagccttcacga.....cagaatggcgatgatcat 501

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 113956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	88.8	17.7	28804	2	US-08-522-874-1
C 2	88.8	17.7	28804	3	US-09-096-942-2
C 3	88.8	17.7	28804	3	US-09-096-867-2
C 4	51.6	10.3	5120	3	US-08-772-270A-6
C 5	51.6	10.3	8370	2	US-08-488-706-1
C 6	51.6	10.3	8370	4	US-09-062-126-1
C 7	48	9.6	1380	4	US-09-252-991A-3550
C 8	48	9.6	1947	4	US-09-252-991A-3557
C 9	47.6	9.5	852	4	US-09-252-991A-11544
C 10	47.6	9.5	1608	4	US-09-252-991A-11665
C 11	47.6	9.5	1608	4	US-09-252-991A-11769
C 12	47	9.4	4403765	3	US-09-103-840A-2
C 13	47	9.4	4411529	3	US-09-103-840A-1
C 14	42	8.4	4403765	3	US-09-103-840A-2
C 15	42	8.4	4411529	3	US-09-103-840A-1
C 16	40.6	8.1	1572	4	US-09-252-991A-3270
C 17	40.6	8.1	1965	4	US-09-252-991A-3234
C 18	40.6	8.1	2325	4	US-09-252-991A-3413
C 19	40.2	8.0	822	4	US-09-252-991A-9029
C 20	40.2	8.0	1527	4	US-09-252-991A-8978
C 21	40.2	8.0	2043	4	US-09-252-991A-8759
C 22	39.2	7.8	2034	4	US-09-252-991A-16314
C 23	39.2	7.8	2190	4	US-09-252-991A-15964
C 24	39	7.8	1254	4	US-09-252-991A-4072
C 25	39	7.8	1404	4	US-09-252-991A-4118
C 26	39	7.8	1680	4	US-09-252-991A-4087
C 27	38.4	7.7	435	4	US-09-252-991A-81

Sequence 105, Appl
Sequence 76, Appl
Sequence 88, Appl
Sequence 206, Appl
Sequence 9, Appl
Sequence 6, Appl
Sequence 1, Appl
Sequence 13089, A
Sequence 12703, A
Sequence 12740, A
Sequence 2823, A
Sequence 12740, A
Sequence 13365, A
Sequence 12592, A
Sequence 13045, A
Sequence 1159, A
Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-592-874-1/C
; Sequence 1, Application US/08592874
; Patent No. 5854034
; GENERAL INFORMATION:
; APPLICANT: POLLOCK, THOMAS J.
; APPLICANT: YAMAZAKI, MOTOHIDE
; APPLICANT: THORNE, LINDA
; APPLICANT: MIKOLAJCZAK, MARCIA
; APPLICANT: ARMENTROUT, RICHARD W.
; TITLE OF INVENTION: DNA SEGMENTS AND METHODS FOR INCREASING
; TITLE OF INVENTION: POLYSACCHARIDE PRODUCTION
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JULES E. GOLDBERG
; STREET: 261 MADISON AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/592.874
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/377,440
; FILING DATE: 24-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDBERG, JULES E.
; REGISTRATION NUMBER: 24,408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-986-4090
; TELEFAX: 212-818-9479
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28804 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (Genomic)
; FRAGMENT TYPE: N-terminal
US-08-592-874-1

Query Match 17.7%; Score 88.8; DB 2; Length 28804;
Best Local Similarity 56.5%; Pred. No. 2.7e-14;

Matches 165; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 80 TTCCTGCCGGCGCCCTCGAAATTCGAGACGCCGCATCTCCACCGGAGACTCACG 139
DB 18808 TTCCTGCCGGCGCCCTCGAGATCATCGAAGTCCGGTGTCCGCCACCGCGCTCAC 18749

QY 140 GCGGCTTGTCTGCTGCTTGTCTTACTCGCGCGTGGGGTGGGGTCTCGGAGATC 199
DB 18748 GCGCGGCTGATGTTGGCGGGCTGGCGATCATCCACCGCTGGCTGGCAATCGCGCGGTG 18689

QY 200 GACATCGTTGCTTCTGCAATCAGAAAGATCGTCCGGCGGACCGTGTAAAGCTGTTTCAG 259
DB 18688 GAAGTGTGGCGCCGACGACGCGCGCATCGCCCGATCGCGGAGACCAAGATCGTCAG 18629

QY 260 CGCTCGAGTGGCGGTGGTGGCGGCGCATCATGTCCGCGATGCCAAACCGTCAAGGCC 319
DB 18628 TCCCGCGAAAGCGGTATCGTCCGGCGCATCTCTGGTGGCGAGGGCGAGAAGTTTCAAG 18569

QY 320 GCGGAGATTCGATCGAGCTGCATCCATTCGCGGTGGTGGATGTTGCGC 371
DB 18568 GCGCAGGTGCTGATCAGCTCGATCCACCATGTCCGAGCGGAGCGCGC 18517

RESULT 2

US-09-096-942-2/c

; Sequence 2, Application US/09096942

; Patent No. 6027925

; GENERAL INFORMATION:

; APPLICANT: Pollock, Thomas J

; APPLICANT: Mikolajczak, Marcia

; APPLICANT: Yamazaki, Motohide

; APPLICANT: Thorne, Linda

; APPLICANT: Armentrout, Richard W

; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria

; FILE REFERENCE: seq list for appl filed from pro. appl

; CURRENT APPLICATION NUMBER: US/09/096,942

; EARLIER FILING DATE: 1998-06-12

; EARLIER FILING DATE: 1997-06-12

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 28804

; TYPE: DNA

; ORGANISM: Sphingomonas sp. S88

US-C9-096-942-2

Query Match 17.7%; Score 88.8; DB 3; Length 28804;

Best Local Similarity 56.5%; Pred. No. 2.7e-14;

Matches 165; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 80 TTCCTGCCGGCGCCCTCGAAATTCGAGACGCCGCATCTCCACCGGAGACTCACG 139
DB 18808 TTCCTGCCGGCGCCCTCGAGATCATCGAAGTCCGGTGTCCGCCACCGCGCTCAC 18749

QY 140 GCGGCTTGTCTGCTGCTTGTCTTACTCGCGCGTGGGGTGGGGTCTCGGAGATC 199
DB 18748 GCGCGGCTGATGTTGGCGGGCTGGCGATCATCCACCGCTGGCTGGCAATCGCGCGGTG 18689

QY 200 GACATCGTTGCTTCTGCAATCAGAAAGATCGTCCGGCGGACCGTGTAAAGCTGTTTCAG 259
DB 18688 GAAGTGTGGCGCCGACGACGCGCGCATCGCCCGATCGCGGAGACCAAGATCGTCAG 18629

QY 260 CGCTCGAGTGGCGGTGGTGGCGGCGCATCATGTCCGCGATGCCAAACCGTCAAGGCC 319
DB 18628 TCCCGCGAAAGCGGTATCGTCCGGCGCATCTCTGGTGGCGAGGGCGAGAAGTTTCAAG 18569

QY 320 GCGGAGATTCGATCGAGCTGCATCCATTCGCGGTGGTGGATGTTGCGC 371
DB 18568 GCGCAGGTGCTGATCAGCTCGATCCACCATGTCCGAGCGGAGCGCGC 18517

RESULT 3

US-09-096-867-2/c

; Sequence 2, Application US/09096867

; Patent No. 6030817

; GENERAL INFORMATION:

; APPLICANT: Pollock, Thomas J

; APPLICANT: Mikolajczak, Marcia

; APPLICANT: Yamazaki, Motohide

; APPLICANT: Thorne, Linda

; APPLICANT: Armentrout, Richard W

; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria

; FILE REFERENCE: seq list for appl filed from pro. appl

; CURRENT APPLICATION NUMBER: US/09/096,867

; EARLIER FILING DATE: 1998-06-11

; EARLIER FILING DATE: 1997-06-12

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 28804

; TYPE: DNA

; ORGANISM: Sphingomonas sp. S88

US-C9-096-867-2

Query Match 17.7%; Score 88.8; DB 3; Length 28804;

Best Local Similarity 56.5%; Pred. No. 2.7e-14;

Matches 165; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 80 TTCCTGCCGGCGCCCTCGAAATTCGAGACGCCGCATCTCCACCGGAGACTCACG 139
DB 18808 TTCCTGCCGGCGCCCTCGAGATCATCGAAGTCCGGTGTCCGCCACCGCGCTCAC 18749

QY 140 GCGGCTTGTCTGCTGCTTGTCTTACTCGCGCGTGGGGTGGGGTCTCGGAGATC 199
DB 18748 GCGCGGCTGATGTTGGCGGGCTGGCGATCATCCACCGCTGGCTGGCAATCGCGCGGTG 18689

QY 200 GACATCGTTGCTTCTGCAATCAGAAAGATCGTCCGGCGGACCGTGTAAAGCTGTTTCAG 259
DB 18688 GAAGTGTGGCGCCGACGACGCGCGCATCGCCCGATCGCGGAGACCAAGATCGTCAG 18629

QY 260 CGCTCGAGTGGCGGTGGTGGCGGCGCATCATGTCCGCGATGCCAAACCGTCAAGGCC 319
DB 18628 TCCCGCGAAAGCGGTATCGTCCGGCGCATCTCTGGTGGCGAGGGCGAGAAGTTTCAAG 18569

QY 320 GCGGAGATTCGATCGAGCTGCATCCATTCGCGGTGGTGGATGTTGCGC 371
DB 18568 GCGCAGGTGCTGATCAGCTCGATCCACCATGTCCGAGCGGAGCGCGC 18517

RESULT 4

US-08-772-270A-6

; Sequence 6, Application US/08772270A

; Patent No. 6019984

; GENERAL INFORMATION:

; APPLICANT: MacInnes, Janet

; APPLICANT: Ricciattti, Paul

; APPLICANT: Mallard, Bonnie

; APPLICANT: Rosendal, Soren

; TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bereskin & Parr

; STREET: 40 King Street West

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5H 3Y2

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM: protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 11:14:15 ; Search time 2864 Seconds

(without alignments)
1771.227 Million cell updates/sec

Title: US-09-484-577A-4

Perfect score: 608

Sequence: 1 LSSLSPRMKSAEVAWGK.....KAGEILIELDPFAGGVGVAT 124

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2889711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09484577/runat_23102003_111410_17988/app_query.fasta_1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITIS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORV=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09484577 @CGN 1 3508 @runat_23102003_111410_17988 -NCFU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPB_CHECK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPEXT=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.ro.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	312	51.3	302650	1	AP005958
C 2	246	40.5	10902	1	AE011858
C 3	243	40.0	11180	1	AE013811
C 4	243	40.0	334050	1	AJ414151
C 5	241	39.6	15552	1	AE004049
C 6	239	39.3	302022	1	AE012558
C 7	238	39.1	1428	1	NME391263
C 8	238	39.1	4254	1	AF121772
C 9	238	39.1	11381	1	AE002524
C 10	238	39.1	349980	6	AY044033
C 11	238	38.1	349980	6	AY044034
C 12	235	38.7	7824	1	NME391260
C 13	235	38.7	326301	1	NMA622491
C 14	231.5	38.1	301708	1	AE016792
C 15	218	35.9	189050	1	AL646077
C 16	215	35.4	28804	1	SSU51197
C 17	215	35.4	28804	6	AR068625
C 18	205	33.7	3526	1	AF273216
C 19	182	29.9	6952	1	AF314503
C 20	182	29.9	6952	1	AF314506
C 21	182	29.9	6952	1	AF314508
C 22	182	29.9	6952	1	AF314509
C 23	182	29.9	6952	1	AF314510
C 24	182	29.9	6952	1	AF314512
C 25	182	29.9	6952	1	AF314514
C 26	182	29.9	6952	1	AF314515
C 27	182	29.9	6952	1	AF414141
C 28	182	29.9	6955	1	AF314511
C 29	182	29.9	6955	1	AF314513
C 30	182	29.9	6958	1	AF314523
C 31	182	29.9	6958	1	AF314524
C 32	182	29.9	6958	1	AF314526
C 33	182	29.9	7742	1	PASLKTICABD
C 34	181	29.8	6469	1	AF314521
C 35	181	29.8	6952	1	AF314517
C 36	181	29.8	6952	1	AF314518
C 37	181	29.8	6952	1	AF314519
C 38	181	29.8	6952	1	AF314520
C 39	181	29.8	6952	1	AF314522
C 40	181	29.8	300883	1	AE016809
C 41	179	29.4	7801	1	PASALIKT
C 42	179	29.4	7841	1	PAJLEUKOT
C 43	177	29.1	7489	1	PAU66586
C 44	177	29.1	7773	1	AF381184
C 45	176	28.9	15746	1	APXIIICAB

ALIGNMENTS

RESULT 1

```

AP005958/c
LOCUS       302650 bp      DNA      linear      BCT 28-MAR-2003
DEFINITION  Bradyrhizobium japonicum USDA 110 DNA, complete genome, section
24/31.
ACCESSION   AP005958
VERSION     AP005958.1
KEYWORDS    GI:27354550
SOURCE      Bradyrhizobium japonicum USDA 110
ORGANISM    Bradyrhizobium japonicum USDA 110
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE   1
AUTHORS     Kaneko,T., Nakamura,Y., Sato,S., Minamisawa,K., Uchiyumi,T.,
            Sasamoto,S., Watanabe,A., Idegawa,K., Iriguchi,M., Kawashima,K.,
            Kohara,M., Matsumoto,M., Shimpo,S., Tsuruoka,H., Wada,T., Yamada,M.
            and Tabata,S.
TITLE       Complete genomic sequence of nitrogen-fixing symbiotic bacterium
            Bradyrhizobium japonicum USDA110
JOURNAL     DNA Res. 9 (6), 189-197 (2002)
MEDLINE     22484998
PUBMED      12597275
REFERENCE   2
AUTHORS     Kaneko,T., Nakamura,Y., Sato,S., Minamisawa,K., Uchiyumi,T.,
            Sasamoto,S., Watanabe,A., Idegawa,K., Iriguchi,M., Kawashima,K.,
            Kohara,M., Matsumoto,M., Shimpo,S., Tsuruoka,H., Wada,T., Yamada,M.
            and Tabata,S.
TITLE       Complete genomic sequence of nitrogen-fixing symbiotic bacterium
            Bradyrhizobium japonicum USDA110 (supplement)
JOURNAL     DNA Res. 9 (6), 225-256 (2002)
MEDLINE     22485002
PUBMED      12597279
REFERENCE   3
AUTHORS     Kaneko,T.
TITLE       Direct Submission
JOURNAL     Submitted (20-NOV-2002) Takakazu Kaneko, Kazusa DNA Research
            Institute, The First Laboratory for Plant Gene Research; 2-6-7
            Kazusa-kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:kaneko@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/rhizobase/,
            Tel:81-438-52-3935 (ex.2338), Fax:81-438-52-3934)
FEATURES             location/Qualifiers
     source           1..302650
                     /organism="Bradyrhizobium japonicum USDA 110"
                     /mol_type="genomic DNA"
                     /strain="USDA110"
                     /db_xref="taxon:224911"
     gene             31..1110
                     /gene="blr6272"
     CDS              31..1110
                     /gene="blr6272"
                     /note="AraC family
                     ORF ID:blr6272"
                     /codon_start=1
                     /transl_table=11
                     /product="transcriptional regulatory protein"
                     /protein_id="BAC51537.1"
                     /db_xref="GI:27354551"
                     /translation="MENLESLSVESFGGLGFLPVNASGRDAKRGAMFTDLQAPAWID
                     LQAEPRDLQAAQAIIPRERGRIRIDRGEAAIEDITISRWEDFRDSSSHATTTPED
                     RYFVGLKATRAKTLDRQVIFDQGTMPAGTLVVSAPKFLAQAQPCAFHLHFISA
                     DHFPAQSPAAGLGLDLRLDPLAAELARALTEHGDAADREFTRCVGQTLAMH:ARLE
                     PPRVNALPKWRLRRVEQYIADKDRICIISILANVAGSRMHFAQFRTATGYEPR
                     EYLLNHRJHAKTLTANTGRPLAEIALAVGFSTQAHFSTVFKRISQSPANRWLASKA
                     EPLAAELPRRRPAARDRWMAAA"
                     complement(1428..3794)
                     /gene="blr6273"
                     /note="AraC family
                     complement(1428..3794)
                     /gene="blr6273"
                     /note="ORF ID:blr6273
                     probable biotin sulfoxide reductase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAC51538.1"

/db_xref="GI:27354552"
/transl_table=11
/translation="MDRYTATMYRNKQMDMDTIGFPDGLDLSDFGKPHSTHWGVS
ARNGAAGLEVRVAVAGDPDNGIINDPFGALRHQARIAQPAIRGWLERGGGDDRRGR
DEFVSWSHEKALDGLDELIRIDTRGPGAVFGSGSAGRFHHAQOVRHLFNIA
MGYVRSVNSYSSGASSVLLPQILAGVEDITKRNVTWEOIATETDILVAFGGMALKNS
MVAGGSIKSHVRSAMBAARKGCEFILVSPUREDLPVEAGAEWMTCTVSTDTALMLG
IVHTLSEKLDQAFLDRTYTGWPFLRYLTVGESDQPKSABWAAIISGVDAITLTL
ARRNLAKRALITVSHLSQRAEHGEQPMGMVLAALGOIGLPGGYAYSLGAIGVYG
RRVNDVPTLGOGRNGVRDIFVARIADMLNPGATYRNGETRTPDIRLVYWAGG
NPFHMHODINLRKAFALDFTFVHELAMTATARHADIVLPSTMTLEREDIGVSNDDP
LMAVHRISAPFGLARDDEIFADLAERLGAPEPFTGRTSREKIKHLIEPTASLAM
KGLAPSFDEFQKGSILVQQDDCGRLRRFPVAAHALPTSPGRIEIFFSTIAGH
GDADPGHPVLMDKTDMPKPGAPCFVLAQNPVTRLHSQDLDFGHSUGAKHMDREVARM
NPVDADARGIKDGDIIIRLFNERGACLAARVHTDGIAPGVQLPTGAWYDMPDPEADP
LCVHGNFNVLTTRDVTSSFAQGTGQLTTVEVEKFTGNLPPIRAFDFV"
complement(3698..4947)
/gene="blr6274"
complement(3698..4947)
/gene="blr6274"
/note="ORF ID:blr6274
putative dehydrogenase"
/codon_start=1
/transl_table=11
/protein_id="BAC51539.1"
/db_xref="GI:27354553"
/translation="MTGASDTVAIGADALRPPFAADLLRAGGFTADQADTAAMLVWAN
LRGADSHGVLRIIPRYVMEVGLINPAAPROVSGKGAIVADIAQAPAGVGMSLAAR
KASLADATHGIGWCAVRSISHAGIYAHQVAAQRIGIVMTASKPLMVIYHSSAEG
VSTNPZTAIAAPTGNPDRPLLMDSTVALGKIMAAADAGQRTPVGAVDVGQVLPDA
PREAKALPMAGPKGSLMIENVLSTANPLIAVALVHGGDPCGNGVLAVLDPDA
FGDNFTLEVDRLGAAIKLPPAAGADSVYLPGRGFEEMERRSRAGIPLARGTVSRLL
ALSKN"
complement(4950..5873)
/gene="blr6275"
complement(4950..5873)
/gene="blr6275"
/note="ORF ID:blr6275
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC51540.1"
/db_xref="GI:27354554"
/translation="MSAPPLSQPTREIPMOEKHTFRSDGLKLSAVLHVDPARKPGQRL
PAFIVLHGFVSGXDESHABIQARMJBQMGYVALRDFRCGSEGERAGVRCFDQVAD
AKNALTFLAEREVDPCRIGVGVSFCAAVSVYAGVDDRFACVJSSCGMGDEKPR
GOHPTPEAAKFTSLLEKGRKOKQETGQSMWISRFDAVPIBEHLRKNLSKALMEIPI
ETAMWYNFRADEVGNIAPRPILFLHTANDTITPTEQSIRMEKAGQAPAEVLITGT
SHFPLAEGDAPRKAIKMWLDKFFFSPLGA"
complement(5910..6812)
/gene="blr6276"
complement(5910..6812)
/gene="blr6276"
/note="ORF ID:blr6276
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAC51541.1"
/db_xref="GI:27354555"
/translation="MPPIHVAHVGHVKINTTAIEAVVRDATEILGLQITVYDEGQWLS
SNGRAEALVLRSHENATHITIGLEALTVDAREESRVSACRVTISRPKSLDCKMAAG
VTEDTPEGLFVHTPTROQIVGRYPVTHGVGNEMDHLNLSPOPIATROQLERICG
LRISEMYDLSLWMSGGRQHHIVGVVGRKVLGHHSFEPKEFNDYKLGGLLDRFD
RMLWNGRHRPGDNTIAYITDSSGAMIECSGSMALIDDDGDFVPVNTNLRPGRNVR
CMNVNGTPPLEWRQFHFPPTAIT"
complement(7080..7853)
/gene="blr6277"
complement(7080..7853)
/gene="blr6277"
/note="GntK family
ORF ID:blr6277"
/codon_start=1
/transl_table=11
/product="transcriptional regulatory protein"

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 11:14:20 ; Search time 1724 Seconds
(without alignments)
1748.118 Million cell updates/sec

Title: US-09-484-577A-4
Perfect score: 608
Sequence: 1 LSLSPRKMSAREVAVGOK.....KAGEILIELDPFAGGVNVAT 124

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09484577/runat_23102003.111410.18001/app_query.fasta_1.263
-DB=EST -Qfmt=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -DOCALIGN=0
-UNITS=bits -START=1 -END=-1 -MATRIX=bloms62 -TRANS=human40.cdi -L-ST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
USER=US09484577 @CEN 1.1 2810 @runat 23102003.111410.18001 -NCPG=6 -ICFJ=3
NO MMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	122	20.1	772	28	AF075872	AF075872 AF075872
C 2	117	19.2	1126	29	B2573090	B2573090 msh2_2939
C 3	105	17.3	750	28	AF075952	AF075952 AF075952
C 4	95.5	15.7	945	10	BE564254	BE564254 601343055
C 5	92	15.1	812	29	B2549462	B2549462 pacsl-60
C 6	87	14.3	961	10	BF781847	BF781847 602107387
C 7	85.5	14.1	455	29	P284R	AL160620 Leishmani
C 8	85	14.0	682	29	AG163349	AG163349 Pan trogl
C 9	84	13.8	432	10	BG274115	BG274115 WHE2231.G
C 10	84	13.8	756	28	AO991716	AO991716 Rfc00179F
C 11	83.5	13.7	610	12	BM631476	BM631476 170006875
C 12	83.5	13.7	973	10	BG758850	BG758850 602713239
C 13	83	13.7	939	29	B2706228	B2706228 PUBLV37T0
C 14	82.5	13.6	431	29	B2883003	B2883003 CH240_241
C 15	82.5	13.6	1175	29	B2558366	B2558366 pacsl-60
C 16	82	13.5	456	12	BJ189099	BJ189099 BJ189099-
C 17	82	13.5	884	29	B2636906	B2636906 OGCCR65TM
C 18	81.5	13.4	563	28	BH775270	BH775270 f2mb011f0
C 19	81.5	13.4	678	14	CD425333	CD425333 SA1.11.H0
C 20	81	13.3	469	9	AU070712	AU070712 AU070712
C 21	81	13.3	906	14	CB575007	CB575007 AGENCOURT
C 22	80	13.2	795	29	CNS05PMK	AL348149 Tetraodon
C 23	80	13.2	1201	9	AJ514080	AL514080 AL514080
C 24	80	13.2	1337	29	B2572952	B2572952 msh2_2874
C 25	79.5	13.1	360	10	BE481913	BE481913 167524.BA
C 26	79.5	13.1	394	14	D27997	D27997 CELK007E1F
C 27	79.5	13.1	423	9	AI969323	AI969323 wq96h04.x
C 28	79.5	13.1	507	13	B0606908	B0606908 BRY_2784
C 29	79.5	13.1	532	13	B0817896	B0817896 1030066H0
C 30	79.5	13.1	768	12	BI909889	BI909889 603070521
C 31	79.5	13.1	1002	13	BX344531	BX344531 BX344531
C 32	79.5	13.1	1088	29	B2577484	B2577484 msh2_5430
C 33	79	13.0	487	10	BG158865	BG158865 RH122_43
C 34	79	13.0	710	12	BJ570919	BJ570919 BJ570919-
C 35	79	13.0	711	12	BJ568037	BJ568037 BJ568037
C 36	79	13.0	946	10	BF305164	BF305164 601892608
C 37	79	13.0	969	10	BF787122	BF787122 602108661
C 38	79	13.0	1159	29	AG071759	AG071759 Pan trogl
C 39	78.5	12.9	499	14	CA621031	CA621031 wlin.pk00
C 40	78.5	12.9	528	14	CB993505	CB993505 AGENCOURT
C 41	78.5	12.9	734	28	AO272038	AO272038 rdbxb00270
C 42	78.5	12.9	775	29	B2577887	B2577887 msh2_5621
C 43	78.5	12.9	825	12	BJ603588	BJ603588 603244491
C 44	78.5	12.9	852	29	B2553772	B2553772 pacsl-60
C 45	78.5	12.9	853	29	CC115510	CC115510 NDL.71C2-

ALIGNMENTS

RESULT 1
AF075872/c
LOCUS AF075872 Salmonella typhimurium LT2, Lambda DASH II Salmonella
DEFINITION typhimurium genomic clone 142-T3, genomic survey sequence.
ACCESSION AF075872
VERSION AF075872.1 GI:3320742
KEYWORDS GSS.
SOURCE Salmonella typhimurium
ORGANISM Salmonella typhimurium
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.
REFERENCE 1 (bases 1 to 772)

AUTHORS Wong, R.M.-Y., Wong, K.K., Benson, N.R., and McClelland, M.
TITLE Sample sequencing of a *Salmonella typhimurium* LT2 lambda library:
 comparison to the *Escherichia coli* K12 genome
JOURNAL FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
MEDLINE 99243757
PUBMED 10227170
COMMENT Contact: McClelland M
 Molecular Biology
 Sidney Kimmel Cancer Center
 3099 Science Park Road, San Diego, CA 92121, USA
 Email: mclelland@lifesci.sdsu.edu
 Class: shotgun.

FEATURES Location/Qualifiers
 1..772
 /organism="Salmonella typhimurium"
 /mol_type="genomic DNA"
 /strain="LT2"
 /db_xref="taxon:602"
 /clone="142-T3"
 /clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
 /note="Vector: Lambda DASH II; sequenced using Li-Cor
 sequencer"
BASE COUNT 169 a 229 c 205 g 167 t 2 others
ORIGIN

Alignment Scores:
 Pred. No.: 0.00187 Length: 772
 Score: 122.00 Matches: 26
 Percent Similarity: 61.11% Conservative: 18
 Best Local Similarity: 36.11% Mismatches: 28
 Query Match: 20.07% Indels: 1
 DB: 28 Gaps: 0
 US-09-484-577A-4 (1-124) x AF075872 (1-772)
 Qy 49 LeuLeuAlaAlaLeuPheTyxCysAlaValAlaTrpAlaGlyLeuGlyArgIleAspIle 68
 Db 336 ATCTGCTACTCTTATTTCTTCATCTCGGCATCTGGCGTGGTGTGGCCAGCTGGATGAG 277
 Qy 69 ValAlaSerAlaSerArgIleValProGlyAspArgValValGlnProLeu 88
 Db 276 GTTTCACCGGACGGGAAGTAGTATCCCACTTCACGCGACAGGTTCTGCAGTGCCTG 217
 Qy 89 GluValGlyValValArgAlaThrHisValArgAspGlyGlnThrValValAlaGlyGlu 108
 Db 216 GATCGCGGCACTTCGGCGAGTTCACGCTGCGGAGCGACAGATTCAGGCTAACCCAG 157
 Qy 109 IleLeuIleGluLeuAspProPheAlaGlyVal 120
 Db 156 ATTGTGCGCGCGCTTGATCCCGAC-GCGTCTGGCGTC 122

RESULT 2
BZ573090/c BZ573090 1126 bp DNA linear GSS 17-DEC-2002
LOCUS msh2_2939.y2 msh pseudomonas aeruginosa genomic clone msh2_2939,
 genomic survey sequence.
DEFINITION
ACCESSION BZ573090.1 GI:27208151
VERSION GSS.
KEYWORDS Pseudomonas aeruginosa
SOURCE Pseudomonas aeruginosa
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1126)
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
 Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol., (2002) In press
COMMENT Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.

FEATURES Location/Qualifiers
 1..1126
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="M5H"
 /db_xref="taxon:287"
 /clone="msh2_2939"
 /clone_lib="msh"
 /notes="Environmental isolate. Whole genomic shotgun
 library."
BASE COUNT 170 a 360 c 314 g 282 t
ORIGIN

Alignment Scores:
 Pred. No.: 0.0107 Length: 1126
 Score: 117.00 Matches: 38
 Percent Similarity: 47.58% Conservative: 21
 Best Local Similarity: 30.65% Mismatches: 51
 Query Match: 19.24% Indels: 14
 DB: 29 Gaps: 5
 US-09-484-577A-4 (1-124) x BZ573090 (1-1126)
 Qy 3 SerLeuSerProArgMetLysSer-----AlaArgGluValValAlaValGlyGlyLys 20
 Db 734 TCATGGACGCGCGCTGAAGCAGTGTGTGAGGGCAACACAGTTTCAGGGTTCGCGACGGCGC 675
 Qy 21 ThrArgAspGlu-----LeuAlaPheLeuProAlaAlaLeuGluIleValGluThr 37
 Db 674 CGAAGGAGCCACCATGATCTCTGACCGCAACCCGCGCGCTGCGCGCAACTGACC 615
 Qy 38 ProProSerProThrAlaArgLeuThrAlaAlaLeuLeuAlaLeuPheTyr----- 55
 Db 614 GACCCGTTGTGACGGCT-----ACCCACCGCGTCTACCGCGCGCTGTCTTGACCCCTG 561
 Qy 56 -----CysAla-----ValAlaTrpAlaGlyLeuGlyArgIleAspIleValAla 70
 Db 560 CTGGTTCGTGCTGTCTTCATCGCTGCGGCGCTGCGCGCAACTGACGAGGTGACC 501
 Qy 71 SerAlaSerArgLysIleValProGlyAspArgValValLeuValGlnProLeuGluVal 90
 Db 500 CGCGGCGACGCTCGGCTCGCTTCAGCGCATCCAGAGATCCAGAGCTTGAGGGGC 441
 Qy 91 GlyValValArgAlaThrHisValArgAspGlyGlnThrValValAlaGlyGluIleLeu 110
 Db 440 GGCATCTCTGACCGCTCTGCTGTGAAGAGCGACCTGCTGGAAGTGGCCGCGCTG 381
 Qy 111 IleGluLeuAsp 114
 Db 380 GTGCGCTCGAC 369

RESULT 3
AF075952 AF075952 750 bp DNA linear GSS 29-AUG-2000
LOCUS typhimurium genomic clone 1256-T3, genomic survey sequence.
DEFINITION
ACCESSION AF075952
VERSION AF075952.1 GI:3320822
KEYWORDS GSS.
SOURCE Salmonella typhimurium
ORGANISM Salmonella typhimurium
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Salmonella.
REFERENCE 1 (bases 1 to 750)
AUTHORS Wong, R.M.-Y., Wong, K.K., Benson, N.R. and McClelland, M.
TITLE Sample sequencing of a *Salmonella typhimurium* LT2 lambda library:
 comparison to the *Escherichia coli* K12 genome
JOURNAL FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
MEDLINE 99243757

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 11:14:14 ; Search time 236 Seconds
(without alignments)
1418.349 Million cell updates/sec

Title: US-09-484-577A-4
Perfect score: 608
Sequence: 1 LSSIPRMSAREVAVGGK.....KAGELLIDPPAGGVAVAT 124

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q/cgn2.1/USPFO.spool/US09484577/runat_23102003.111409.17977/app_query.fasta_1.263
-DB=N_Geneseq_19jun03 -QMT=rascap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09484577 @CGN 1.1.0 @runat_23102003.111409.17977 -NCPU=6 -ICPU=3
-NO MMAP -LARGUEURY -NEG SCORES=0 -WAIT -DISPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_19jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	238	39.1	1428	21	AAZ54335	Neisseria meningit
2	238	39.1	24158	21	AA81532	N. meningitidis pa
C 3	238	39.1	34980	21	AAF21611	Neisseria meningit
C 4	238	39.1	34980	21	AAF21612	Neisseria meningit
C 5	238	39.1	837096	21	AAA81489	N. meningitidis pa
C 6	235	38.7	1305	21	AAZ54336	Neisseria meningit
C 7	215	35.4	28804	17	AAT37329	Sphingomonas genus
C 8	215	35.4	28804	18	AAT92474	Sphingomonas S88 s
C 9	215	35.4	28804	20	AAV99812	Chromosomal fragme
C 10	215	35.4	28804	20	AAV81474	pkx gene of Pasteu
11	190	31.2	7184	15	AAQ0050	Sequence encoding
12	182	29.9	7183	11	AAQ06074	E. coli CFT073 gen
13	176	28.9	76804	24	ABS78942	ApixBD gene. Acti
14	169	27.8	5120	18	AAT73218	A. pleuropneumonia
15	169	27.8	5120	21	AAZ88585	A. pleuropneumonia
16	168	27.6	7721	18	AAT73220	ApixIABCD gene.
17	168	27.6	7721	21	AAZ88587	A. pleuropneumonia
C 18	154	25.3	3576	19	AAV31265	E. coli J96 pathog
19	124.5	20.5	1245	22	ABA8180	Escherichia coli p
20	124.5	20.5	48012	24	ABS79084	E. coli CFT073 gen
21	124.5	20.5	48254	22	ABA89141	Escherichia coli p
22	124.5	20.5	48345	22	ABA89142	Escherichia coli p
23	104	17.1	6400	20	AAZ22701	Pseudomonas fluore
24	99	16.3	4466	17	AAT39662	Esterase secretory
C 25	91	15.0	2382	24	AAZ94828	Human DNA sequence
C 26	90.5	14.9	9320	22	AAH45588	DNA encoding hydro
27	90	14.8	1302	23	ABL41686	ABC transporter fl
28	90	14.8	8580	23	ABL41682	ABC transporter ge
29	89	14.6	5660	20	AAZ22700	Pseudomonas fluore
C 30	88	14.5	834	23	ABL26523	Drosophila melanog
C 31	88	14.5	2834	23	ABL26522	Drosophila melanog
C 32	88	14.5	13518	20	AAZ20563	Polynucleotide seq
C 33	87.5	14.4	4403765	22	AAI99683	Mycobacterium tube
C 34	87.5	14.4	4411529	22	AAI99682	Mycobacterium tube
C 35	86	14.1	11508	23	AAZ54335	Propionibacterium
36	81.5	13.4	1395	21	AAAG5929	E. coli proliferat
37	81.5	13.4	1395	22	AAH84642	E. coli growth and
38	81.5	13.4	2310	23	AAZ82449	DNA encoding novel
39	80.5	13.2	3462	22	AAZ501509	Bacillus subtilis
C 40	80.5	13.2	4540	24	ABK63777	Rat sequence diffe
C 41	79	13.0	19053	21	AAZ59146	Nucleotide sequenc
C 42	78.5	12.9	357	21	AAZ75328	Human ORFX ORF883
43	78.5	12.9	357	24	AAZ21772	Human ORFX polynuc
C 44	78	12.8	6848	22	AAK83993	Human immune/haema
C 45	77.5	12.7	2556	25	ABZ40207	N. gonorrhoeae nuc

ALIGNMENTS

RESULT 1

AAZ54335

ID AAZ54335 standard; DNA; 1428 BP.

XX

AC AAZ54335;

XX

DT 21-MAR-2000 (first entry)

XX

DE Neisseria meningitidis ORF 764 partial DNA sequence SEQ ID NO:2619.

XX

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
antibacterial; gene therapy; ds.

XX

OS Neisseria meningitidis.

XX

PN W09957280-A2.


```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:      October 23, 2003, 12:11:35 ; Search time 234 Seconds
           (without alignments)
           1421.125 Million cell updates/sec

Title:       US-09-484-577A-4
Perfect score: 608
Sequence:    1 LSLSPKMSAREVVAVGK.....KAGEILIELDPFAGGVGVAT 124

Scoring table: BLOSUM62
              Xgapop 10.0 , Xgapext 0.5
              Ygapop 10.0 , Ygapext 0.5
              Fgapop  6.0 , Fgapext 7.0
              Delop   6.0 , Delext  7.0

Searched:    1792395 seqs, 1340900451 residues
Total number of hits satisfying chosen parameters: 3584790

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Command line parameters:
MODEJ=frame_p2n.model -DEV=xlh
Q=/cgn2_1/USFT0_spool/US09484577/runat_23102003_111412_18082/app_query.fasta.1.263
DB=Published Applications NA -QMT=fastap -SUFFIX=rpnb -MINMATCH=0.1
LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
TRANS=human40.cgi -LIST=45 -LOCAL=LOCAL -OUTFMT=pct -THR SCORE=pct -THR MAX=100
THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HGAPSIZE=500 -MINLEN=0
MAXLEN=200000000 -USER=US09484577 @CGN 1.1.285 @runat 23102003_111412_18082
NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -18082
LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase : Published Applications NA:
1: /cgn2_6/ptodata/1/pubna/us07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/pct_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES
      Query
result No. Score Match Length DB ID Description
-----
1 242 39.8 1428 11 US-09-884-696-36 Sequence 36, Appl

```

```

Sequence 109, Appl
Sequence 79, Appl
Sequence 906, Appl
Sequence 251, Appl
Sequence 829, Appl
Sequence 9, Appl
Sequence 39, Appl
Sequence 83, Appl
Sequence 6704, Appl
Sequence 121, Appl
Sequence 270, Appl
Sequence 1684, Appl
Sequence 260, Appl
Sequence 3629, Appl
Sequence 1, Appl
Sequence 3641, Appl
Sequence 145, Appl
Sequence 6676, Appl
Sequence 97, Appl
Sequence 120, Appl
Sequence 87, Appl
Sequence 4268, Appl
Sequence 4611, Appl
Sequence 2245, Appl
Sequence 1336, Appl
Sequence 135196,
Sequence 649, Appl
Sequence 2093, Appl
Sequence 9794, Appl
Sequence 3389, Appl
Sequence 105, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 4119, Appl
Sequence 13, Appl
Sequence 6015, Appl
Sequence 4, Appl
Sequence 8263, Appl
Sequence 4413, Appl
Sequence 324, Appl
Sequence 6586, Appl

ALIGNMENTS

RESULT 1
US-09-884-696-36
; Sequence 36, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
; APPLICANT: GEORGE, LISLE W
; APPLICANT: ANGELOS, JOHN A
; APPLICANT: HESS, JOHN F
; TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, ANTIBODIES
; TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
; TITLE OF INVENTION: BOVIS INFECTIONS
; FILE REFERENCE: 481.06
; CURRENT APPLICATION NUMBER: US/09/884,696
; CURRENT FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; TYPE: DNA
; LENGTH: 1428
; ORGANISM: Moraxella bovis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1425)
US-09-884-696-36

```

```

Alignment Scores:
Pred. No.: 2,37e-23 Length: 1428
Score: 242.00 Matches: 50
Percent Similarity: 65.79% Conservative: 25
Best Local Similarity: 43.86% Mismatches: 39
Query Match: 39.80% Indels: 0
DB: 11 Gaps: 0

US-09-484-577A-4 (1-124) x US-09-884-696-36 (1-1428)

Qy 10 SerAlaArgGluValValAlaValAlaGlyGlyLysThrArgAspGluLeuAlaPheLeuPro 29
Db 64 GCAGTTCGAGCCACCACTAACCCCTCTTAAGCTACTAAAGAGAACTCGCTTTCTTCT 123
Qy 30 AlaAlaLeuGluLeuValGluThrProSerProThrAlaArgLeuThrAlaAlaLeu 49
Db 124 GCACATCTAGAACTCACTGACACACCTGATCCAGATCTTCTAAGTGGACACCTAGATA 183
Qy 50 LeuAlaAlaLeuPheTyrcysAlaValAlaAlaTTPAlaGlyLeuGlyArgileAspIleVal 69
Db 184 ATCATGATATTGTCCTATTGCTTTGCTATGCTCTGGTTGGACAGATTGACATTTGTT 243
Qy 70 AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValGlnProLeuGlu 89
Db 244 GCTACAGCTTCAGGTAAATTTCTTCAGGTAGCGTAGCAGACACTATTCAATCTTTGGA 303
Qy 90 ValGlyValValAlaArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyGluLe 109
Db 304 ACAGCGATAGTTAAAGCAGTTTATGTACGTGATGGTCAAAATGTTCAACAAGGTGAAATA 363
Qy 110 LeuIleGluLeuAspProPheAlaGlyGlyValAspValAla 123
Db 364 TTAGTAGATTAGTGGGAATCGGTTCCAGATAGTGAATGTTGCT 405

RESULT 2
US-10-085-959-109
; Sequence 109, Application US/10085959
; Publication No. US20030165870A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; APPLICANT: Welch, Rodney A.
; APPLICANT: Burland, Valerie D.
; TITLE OF INVENTION: No. US20030165870A1e1 Sequence of E. Coli CFT073
; FILE REFERENCE: 960296.97648
; CURRENT APPLICATION NUMBER: US/10/085,959
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/242,412
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 255
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 109
; LENGTH: 76804
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (39751)..(39751)
; OTHER INFORMATION: Unsure
; NAME/KEY: misc feature
; LOCATION: (49372)..(49372)
; OTHER INFORMATION: Unsure
; NAME/KEY: misc feature
; LOCATION: (66539)..(66539)
; OTHER INFORMATION: Unsure
; NAME/KEY: misc feature
; LOCATION: (73693)..(73693)
; OTHER INFORMATION: Unsure
; NAME/KEY: misc feature
; LOCATION: (73705)..(73705)
; OTHER INFORMATION: Unsure
; NAME/KEY: misc feature
; LOCATION: (73737)..(73737)
; OTHER INFORMATION: Unsure

NAME/KEY: misc feature
LOCATION: (73759)..(73759)
OTHER INFORMATION: Unsure
NAME/KEY: misc feature
LOCATION: (73761)..(73761)
OTHER INFORMATION: Unsure
NAME/KEY: misc feature
LOCATION: (73766)..(73766)
OTHER INFORMATION: Unsure
NAME/KEY: misc feature
LOCATION: (73774)..(73774)
OTHER INFORMATION: Unsure
NAME/KEY: misc feature
LOCATION: (73781)..(73781)
OTHER INFORMATION: Unsure
NAME/KEY: misc feature
LOCATION: (73798)..(73798)
OTHER INFORMATION: Unsure
NAME/KEY: misc feature
LOCATION: (73802)..(73802)
OTHER INFORMATION: Unsure
US-10-085-959-109
Alignment Scores:
Pred. No.: 6.11e-12 Length: 76804
Score: 176.00 Matches: 36
Percent Similarity: 60.78% Conservative: 26
Best Local Similarity: 35.29% Mismatches: 40
Query Match: 28.95% Indels: 0
DB: 12 Gaps: 0

US-09-484-577A-4 (1-124) x US-10-085-959-109 (1-76804)

Qy 20 LysThrArgAspGluLeuAlaPheLeuProAlaAlaLeuGluLeuValGluThrProPro 39
Db 17983 CGTAAAGACGCAAAATGAATTTCTACCGCTCATCTGGAATTAATTAAGAACGCGGTA 18042
Qy 40 SerProThrAlaArgLeuThrAlaAlaLeuLeuAlaLeuPheTyrcysAlaValAla 59
Db 18043 TCAGACGCCCGCTCTGTTGCTTATTTATTTATGGGGTTTCTGTTATGCTGCTATT 18102
Qy 60 TrpAlaGlyLeuGlyArgileAspIleValAlaSerAlaSerArgLysIleValProGly 79
Db 18103 TTATCTGTTTATAGTCAGGTGGAATTTGTCGCTGCAAAATGGAATTAACACTAAGT 18162
Qy 80 AspArgValLysLeuValGlnProLeuGluValAlaGlyValValAlaThrHisValArg 99
Db 18163 GCGCGCAGCAAGAATAATTAACCTATTGAAACTCAATAGTTAAAGAAATTAATTCGTA 18222
Qy 100 AspGlyGlnThrValLysAlaGlyGluLeuLeuLeuGluLeuAspProPheAlaGly 119
Db 18223 GAAGGAGAGTCAGTCGGAAGCGGATGTTATTAAAGCTTACAGCACTGGAGCTGAA 18282
Qy 120 valasp 121
Db 18283 GCTGAT 18288

RESULT 3
US-09-956-004-79/c
; Sequence 79, Application US/09956004
; Patent No. US20020072595A1
; GENERAL INFORMATION:
; APPLICANT: Patrick J. Dillon et al.
; TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands
; FILE REFERENCE: PB324D1
; CURRENT APPLICATION NUMBER: US/09/956,004
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 08/976,259
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/061,953
; PRIOR FILING DATE: 1997-10-14
; PRIOR APPLICATION NUMBER: 60/031,626
; PRIOR FILING DATE: 1996-11-22

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 11:14:20 ; Search time 57 Seconds
(without alignments)
960.202 Million cell updates/sec

Title: US-09-484-577A-4

Perfect score: 608
Sequence: 1 LSSLSPRMKSREVAVGKK.....KAGBILIELDPFAGGVVAT 124

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame+p2n model -DEV=xlh

Q=/cgn2_1/USPTO_spool/US09484577/runat 23102003_111411_18018/app_query.fasta_1.263

DB=Issued_Patents_NA -Qfmr=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

LOOPEXT=0 -UNITS=BITS -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi

LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15

MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000

USER=US09484577 @CGN 1.1.56 @runat 23102003_111411_18018 -NCPU=6 -ICPU=3

NO MWAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

DSV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	ID	Description
C 1	215	35.4	28804	2	US-08-592-874-1 Sequence 1, Appli
C 2	215	35.4	28804	3	US-09-096-942-2 Sequence 2, Appli
C 3	215	35.4	28804	3	US-09-096-867-2 Sequence 2, Appli
C 4	169	27.8	5120	3	US-08-772-270A-6 Sequence 6, Appli
C 5	169	27.8	8370	2	US-08-488-706-1 Sequence 1, Appli
C 6	169	27.8	8370	4	US-09-062-126-1 Sequence 1, Appli
C 7	168	27.6	7721	3	US-08-772-270A-14 Sequence 14, Appli
C 8	154	25.3	3576	4	US-08-976-259-79 Sequence 79, Appli
C 9	117	19.2	1236	4	US-09-328-352-2593 Sequence 2593, Ap
C 10	109	17.9	1263	4	US-09-252-991A-13467 Sequence 13467, A
C 11	109	17.9	1263	4	US-09-252-991A-13658 Sequence 13658, A
C 12	104.5	17.2	1360	4	US-09-252-991A-3550 Sequence 3550, Ap

13	104.5	17.2	1947	4	US-09-252-991A-3557 Sequence 3557, Ap
C 14	104	17.1	852	4	US-09-252-991A-11544 Sequence 11544, A
C 15	104	17.1	1608	4	US-09-252-991A-11665 Sequence 11665, A
C 16	104	17.1	1608	4	US-09-252-991A-11769 Sequence 11769, A
C 17	103	16.9	25165	4	US-09-453-702B-39 Sequence 39, Appli
C 18	99	16.3	4465	2	US-08-620-605D-1 Sequence 1, Appli
C 19	99	16.3	4547	2	US-09-005-232A-1 Sequence 1, Appli
C 20	90	14.8	1302	4	US-09-491-772-5 Sequence 5, Appli
C 21	90	14.8	8580	4	US-09-491-772-1 Sequence 1, Appli
C 22	87.5	14.4	4403765	3	US-09-103-840A-2 Sequence 2, Appli
C 23	87.5	14.4	4411529	3	US-09-103-840A-1 Sequence 1, Appli
C 24	82	13.5	885	4	US-09-252-991A-11189 Sequence 11189, A
C 25	82	13.5	1245	4	US-09-252-991A-11257 Sequence 11257, A
C 26	82	13.5	1386	4	US-09-252-991A-10994 Sequence 10994, A
C 27	80	13.2	1011	4	US-09-328-352-4076 Sequence 4076, Ap
C 28	79.5	13.1	945	4	US-09-252-991A-3765 Sequence 3765, Ap
C 29	79.5	13.1	1383	4	US-09-252-991A-3609 Sequence 3609, Ap
C 30	79.5	13.1	1533	4	US-09-252-991A-8687 Sequence 8687, Ap
C 31	78	12.8	2398	4	US-09-453-702B-260 Sequence 260, App
C 32	77	12.7	49272	1	US-08-614-770A-1 Sequence 1, Appli
C 33	76	12.5	1251	4	US-09-252-991A-444 Sequence 444, App
C 34	76	12.5	1362	4	US-09-252-991A-417 Sequence 417, App
C 35	76	12.5	2256	4	US-09-252-991A-16502 Sequence 16502, A
C 36	75.5	12.4	1321	4	US-09-175-658B-20 Sequence 20, Appl
C 37	75.5	12.4	1543	4	US-09-364-230-17 Sequence 17, Appl
C 38	74.5	12.3	3603	1	US-08-188-582-15 Sequence 15, Appl
C 39	74.5	12.3	3603	1	US-08-646-715-15 Sequence 15, Appl
C 40	73	12.0	399	4	US-09-252-991A-14158 Sequence 14158, A
C 41	73	12.0	1161	4	US-09-252-991A-14117 Sequence 14117, A
C 42	73	12.0	1233	4	US-09-252-991A-14019 Sequence 14019, A
C 43	73	12.0	1239	4	US-09-252-991A-8514 Sequence 8514, Ap
C 44	73	12.0	1788	4	US-09-252-991A-10460 Sequence 10460, A
C 45	73	12.0	1911	4	US-09-252-991A-8331 Sequence 8331, Ap

ALIGNMENTS

US-08-592-874-1/c
; Sequence 1, Application US/08592874
; Patent No. 5854034
; GENERAL INFORMATION:
; APPLICANT: POLLOCK, THOMAS J.
; APPLICANT: YAMAZAKI, MOTOHIDE
; APPLICANT: THORNE, LINDA
; APPLICANT: MIKOLAJCZAK, MARCIA
; APPLICANT: ARMENTROUT, RICHARD W
; TITLE OF INVENTION: DNA SEGMENTS AND METHODS FOR INCREASING
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JULES E. GOLDBERG
; STREET: 261 MADISON AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592.874
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/377,440
; FILING DATE: 24-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDBERG, JULES E.
; REGISTRATION NUMBER: 24,408
; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 212-986-4090
; TELEFAX: 212-818-9479
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28804 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FRAGMENT TYPE: N-terminal
; US-08-592-874-1

Alignment Scores:
Pred. No.: 1.35e-15 Length: 28804
Score: 215.00 Matches: 44
Percent Similarity: 60.82% Conservative: 15
Best Local Similarity: 45.36% Mismatches: 38
Query Match: 35.36% Indels: 0
DB: 2 Gaps: 0

US-09-484-577A-4 (1-124) x US-08-592-874-1 (1-28804)
QY 27 PheLeuProAlaAlaLeuGluIleValGluThrProSerProThrAlaArgLeuThr 46
Db 18808 TTCTGCGCGCGCGCTCGAGATCATCGAAGCTCCGGTGTCCGCCACCGCGGCTCACC 18749
QY 47 AlaAlaLeuAlaAlaLeuPheTyrCysAlaValAlaTrpAlaGlyLeuGlyArgIle 66
Db 18748 GGCCGGGTGATGTGGCGGGCTGGCGATCACCACCGCTGGTGGCAATCGCGCGGTG 18689
QY 67 AspIleValAlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValGln 86
Db 18688 GAAGTGTGGCGCGCGACGCGAGGCGCGATCGCCCGATCGCGAGACCAAGATCGTGCGAG 18629
QY 87 ProLeuGluValGlyValValArgAlaThrHisValArgAspGlyGlnThrValLysAla 106
Db 18628 TCCCGCGAAGCGGTATCGTCCGCGCATTCGTGGTGGCGAGGGGCGAGAGGTTCAAGAG 18569
QY 107 GlyGluIleLeuIleGluLeuAspProPheAlaGlyValValAspValAla 123
Db 18568 GGCCAGGTGCTGATCAGCTCGATCCACCATGTCCGCGACCGCGAGCGCGG 18518

RESULT 3
US-09-096-867-2/c
; Sequence 2, Application US/09096867
; Patent No. 6030817
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for Xanthomonas Campestris
; CURRENT APPLICATION NUMBER: US/09/096.867
; CURRENT FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 28804
; TYPE: DNA
; ORGANISM: Sphingomonas sp. S88
US-09-096-867-2

Alignment Scores:
Pred. No.: 1.35e-15 Length: 28804
Score: 215.00 Matches: 44
Percent Similarity: 60.82% Conservative: 15
Best Local Similarity: 45.36% Mismatches: 38
Query Match: 35.36% Indels: 0
DB: 3 Gaps: 0

US-09-484-577A-4 (1-124) x US-09-096-867-2 (1-28804)
QY 27 PheLeuProAlaAlaLeuGluIleValGluThrProSerProThrAlaArgLeuThr 46
Db 18808 TTCTGCGCGCGCGCTCGAGATCATCGAAGCTCCGGTGTCCGCCACCGCGGCTCACC 18749
QY 47 AlaAlaLeuAlaAlaLeuPheTyrCysAlaValAlaTrpAlaGlyLeuGlyArgIle 66
Db 18748 GGCCGGGTGATGTGGCGGGCTGGCGATCACCACCGCTGGTGGCAATCGCGCGGTG 18689
QY 67 AspIleValAlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValGln 86
Db 18688 GAAGTGTGGCGCGCGACGCGAGGCGCGATCGCCCGATCGCGAGACCAAGATCGTGCGAG 18629
QY 87 ProLeuGluValGlyValValArgAlaThrHisValArgAspGlyGlnThrValLysAla 106
Db 18628 TCCCGCGAAGCGGTATCGTCCGCGCATTCGTGGTGGCGAGGGGCGAGAGGTTCAAGAG 18569
QY 107 GlyGluIleLeuIleGluLeuAspProPheAlaGlyValValAspValAla 123
Db 18568 GGCCAGGTGCTGATCAGCTCGATCCACCATGTCCGCGACCGCGAGCGCGG 18518

RESULT 2
US-09-096-942-2/c
; Sequence 2, Application US/09096942
; Patent No. 6027925
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for Xanthomonas Campestris
; CURRENT APPLICATION NUMBER: US/09/096.942
; CURRENT FILING DATE: 1998-06-12
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 28804
; TYPE: DNA
; ORGANISM: Sphingomonas sp. S88
US-09-096-942-2

Alignment Scores:
Pred. No.: 1.35e-15 Length: 28804
Score: 215.00 Matches: 44
Percent Similarity: 60.82% Conservative: 15
Best Local Similarity: 45.36% Mismatches: 38
Query Match: 35.36% Indels: 0

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 14:46:07 ; Search time 2169 Seconds
(without alignments)
9449.388 Million cell updates/sec

Title: US-09-484-577A-3

Perfect score: 501

Sequence: 1 actctcagcctctcaccga.....cagaatgcggcgatgatcat 501

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2016740

Minimum DB seq length: 0

Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	24	4.8	60	6	AR231929	Sequence AR231929
2	24	4.8	60	6	AR231931	Sequence AR231931
3	24	4.8	60	6	AR231931	Sequence AR231931
4	24	4.8	60	6	AR231931	Sequence AR231931
5	24	4.8	60	6	AX464727	Sequence AX464727
6	24	4.8	60	6	AX464727	Sequence AX464727
7	24	4.8	60	6	AX464729	Sequence AX464729
8	24	4.8	60	6	AX464729	Sequence AX464729
9	21	4.2	24	6	A68884	Sequence A68884
10	21	4.2	24	6	AR055096	Sequence AR055096
11	21	4.2	24	6	AR055098	Sequence AR055098
12	21	4.2	24	6	AR068437	Sequence AR068437
13	21	4.2	24	6	AR068439	Sequence AR068439
14	21	4.2	24	6	AR121280	Sequence AR121280
15	21	4.2	24	6	AR231914	Sequence AR231914
16	21	4.2	24	6	AX464709	Sequence AX464709
17	21	4.2	24	6	BD062984	DNA and s BD062984
18	21	4.2	24	6	I13333	Sequence I13333
19	21	4.2	24	6	I19026	Sequence I19026
20	21	4.2	28	6	AR035013	Sequence AR035013
21	21	4.2	28	6	AR082097	Sequence AR082097
22	21	4.2	28	6	AR118488	Sequence AR118488
23	21	4.2	28	6	AR151056	Sequence AR151056
24	21	4.2	28	6	AR218212	Sequence AR218212
25	21	4.2	28	6	AR222652	Sequence AR222652
26	21	4.2	28	6	AR231690	Sequence AR231690
27	20	4.0	23	6	AX194017	Sequence AX194017
28	20	4.0	24	6	A68887	Sequence A68887
29	20	4.0	24	6	AR034959	Sequence AR034959
30	20	4.0	24	6	AR034973	Sequence AR034973
31	20	4.0	24	6	AR082043	Sequence AR082043
32	20	4.0	24	6	AR082057	Sequence AR082057
33	20	4.0	24	6	AR082102	Sequence AR082102
34	20	4.0	24	6	AR082104	Sequence AR082104
35	20	4.0	24	6	AR100939	Sequence AR100939
36	20	4.0	24	6	AR100941	Sequence AR100941
37	20	4.0	24	6	AR100943	Sequence AR100943
38	20	4.0	24	6	AR100945	Sequence AR100945
39	20	4.0	24	6	AR100965	Sequence AR100965
40	20	4.0	24	6	AR100979	Sequence AR100979
41	20	4.0	24	6	AR101019	Sequence AR101019
42	20	4.0	24	6	AR101021	Sequence AR101021
43	20	4.0	24	6	AR118434	Sequence AR118434
44	20	4.0	24	6	AR118448	Sequence AR118448
45	20	4.0	24	6	AR130747	Sequence AR130747
46	20	4.0	24	6	AR151002	Sequence AR151002
47	20	4.0	24	6	AR151016	Sequence AR151016
48	20	4.0	24	6	AR199145	Sequence AR199145
49	20	4.0	24	6	AR199149	Sequence AR199149
50	20	4.0	24	6	AR199163	Sequence AR199163
51	20	4.0	24	6	AR218158	Sequence AR218158
52	20	4.0	24	6	AR218172	Sequence AR218172
53	20	4.0	24	6	AR222598	Sequence AR222598
54	20	4.0	24	6	AR222612	Sequence AR222612
55	20	4.0	24	6	AR231636	Sequence AR231636
56	20	4.0	24	6	AR231650	Sequence AR231650
57	20	4.0	24	6	BD062987	Sequence BD062987
58	20	4.0	24	6	BD080944	DNA and s BD080944
59	20	4.0	48	6	AR035015	Sequence AR035015
60	20	4.0	48	6	AR082099	Sequence AR082099
61	20	4.0	48	6	AR118490	Sequence AR118490
62	20	4.0	48	6	AR151058	Sequence AR151058
63	20	4.0	48	6	AR218214	Sequence AR218214
64	20	4.0	48	6	AR222654	Sequence AR222654
65	20	4.0	48	6	AR231692	Sequence AR231692

Pred. No. is the number of results predicted by chance to have a

66	20	4.0	53	6	AR035017	AR035017	Sequence
67	20	4.0	53	6	AR082101	AR082101	Sequence
68	20	4.0	53	6	AR118492	AR118492	Sequence
69	20	4.0	53	6	AR151060	AR151060	Sequence
70	20	4.0	53	6	AR218216	AR218216	Sequence
71	20	4.0	53	6	AR223656	AR223656	Sequence
72	20	4.0	53	6	AR231694	AR231694	Sequence
73	19	3.8	24	6	AR121274	AR121274	Sequence
74	19	3.8	24	6	AR139761	AR139761	Sequence
75	19	3.8	24	6	AR165220	AR165220	Sequence
76	19	3.8	24	6	AR201391	AR201391	Sequence
77	19	3.8	24	6	AR208439	AR208439	Sequence
78	19	3.8	24	6	AR225774	AR225774	Sequence
79	19	3.8	24	6	AR230325	AR230325	Sequence
80	19	3.8	24	6	AR243315	AR243315	Sequence
81	19	3.8	24	6	AR264180	AR264180	Sequence
82	19	3.8	24	6	AR279223	AR279223	Sequence
83	19	3.8	24	6	AR310020	AR310020	Sequence
84	19	3.8	24	6	AX023539	AX023539	Sequence
85	19	3.8	24	6	AX301748	AX301748	Sequence
86	19	3.8	24	6	AX347891	AX347891	Sequence
87	19	3.8	24	6	AX348098	AX348098	Sequence
88	19	3.8	24	6	AX384667	AX384667	Sequence
89	19	3.8	24	6	AX459947	AX459947	Sequence
90	19	3.8	24	6	AX470041	AX470041	Sequence
91	19	3.8	24	6	BD016439	BD016439	Gene der1
92	19	3.8	24	6	BD080835	BD080835	Asthma-re
93	19	3.8	24	6	BD082003	BD082003	Asthma as
94	19	3.8	24	6	BD086207	BD086207	Transcrip
95	19	3.8	24	6	BD106876	BD106876	Isclared
96	19	3.8	24	6	113327	113327	Sequence 1
97	19	3.8	24	6	119020	119020	Sequence 1
98	19	3.8	24	6	127658	127658	Sequence 13
99	19	3.8	24	6	127653	127653	Sequence 8
100	19	3.8	48	6	127649	127649	Sequence 4

ALIGNMENTS

```

RESULT 1
LOCUS AR231929 PAT 20-DEC-2002
DEFINITION Sequence 19 from patent US 6455255. linear
ACCESSION AR231929
VERSION AR231929.1 GI:27273523
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 60)
Birkenmeyer,L.G., Leary,T.P., Muerhoff,A.S., Desai,S.M. and
Mushahwar,I.K.
TITLE Method of performing subtractive hybridization using RDA
JOURNAL Patent: US 6455255-A 1924-SEP-2002;
FEATURES Location/Qualifiers
source 1..60
/organism="unknown"
BASE COUNT 11 a 17 c 17 g 11 t 4 others
ORIGIN

Query Match 4.8%; Score 24; DB:6; Length 6C;
Best Local Similarity 100.0%; Pred.No. 0.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ACTCTCCAGCCTCTCACCGAGGAT 24
|||||
Db 4 ACTCTCCAGCCTCTCACCGAGGAT 27
|||||

RESULT 2
LOCUS AR231929/c PAT 20-DEC-2002

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 15:26:52 ; Search time 1679 Seconds
(without alignments)
7252.259 Million cell updates/sec

Title: US-09-484-577A-3

Perfect score: 501

Sequence: 1 actctccagctctccaccga.....cagaatgcgcatgatcat 501

Scoring table: OLIGO_NUC

Gapop_60.0 , Gapext 60.0

Searched: 2278132 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4672214

Minimum DB seq length: 0

Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_plt:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rdc:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gsl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	4.8	131	9	AF091519
C 2	22	4.4	104	14	CA453170
C 3	20	4.0	178	28	AZ924992 Bt44 Homo
C 4	19	3.8	82	10	BG599357

C	5	19	3.8	100	10	BG599358
	6	19	3.8	109	10	BG315516
	7	19	3.8	121	10	BG315546
	8	19	3.8	131	10	BG315530
	9	19	3.8	132	10	BE644247
	10	19	3.8	189	9	AF027371
	11	19	3.8	214	9	AF027372
	12	19	3.8	219	9	AF027370
	13	19	3.8	229	9	AF027374
	14	19	3.8	241	9	AJ414389
	15	19	3.8	243	9	AJ414388
C	16	18	3.6	138	9	AF091520
C	17	18	3.6	224	12	BG946550
	18	17	3.4	103	9	AA234634
	19	17	3.4	107	28	BH216229
	20	17	3.4	127	10	BE588545
	21	17	3.4	163	9	AI192272
C	22	17	3.4	167	12	BM706517
	23	17	3.4	194	14	T59467
	24	17	3.4	194	12	BM719924
	25	17	3.4	216	14	N90915
	26	17	3.4	227	9	AI910087
C	27	17	3.4	232	14	CB933844
	28	17	3.4	235	9	AA841745
	29	17	3.4	236	14	CB899983
	30	17	3.4	240	9	AI561896
C	31	17	3.4	244	10	BG360850
C	32	17	3.4	247	14	CB031588
	33	16	3.2	87	28	AZ594197
	34	16	3.2	104	12	B1396353
C	35	16	3.2	110	10	BG354695
	36	16	3.2	114	13	BM300094
C	37	16	3.2	117	28	AZ893616
C	38	16	3.2	119	29	TA209G05P
C	39	16	3.2	135	9	AA463183
	40	16	3.2	141	12	BJ234310
	41	16	3.2	144	10	BE497583
	42	16	3.2	149	29	CC178168
	43	16	3.2	159	14	CB066573
	44	16	3.2	160	10	BE497606
	45	16	3.2	161	13	BY008185
	46	16	3.2	166	12	BY111777
C	47	16	3.2	168	9	AV414210
	48	16	3.2	169	9	AL121138
C	49	16	3.2	169	12	BM174574
C	50	16	3.2	171	29	BM174574
C	51	16	3.2	171	29	BM174574
	52	16	3.2	179	10	BE438922
	53	16	3.2	179	14	CA347516
C	54	16	3.2	180	28	BH863430
C	55	16	3.2	181	14	CB473406
C	56	16	3.2	189	9	AF027371
C	57	16	3.2	192	28	AZ922877
C	58	16	3.2	194	9	AI903647
	59	16	3.2	196	9	AA066211
	60	16	3.2	196	9	AA931123
	61	16	3.2	198	13	BQ401278
	62	16	3.2	198	14	CA698323
C	63	16	3.2	202	14	CA659563
C	64	16	3.2	206	29	FR0011531
C	65	16	3.2	214	12	BJ322497
C	66	16	3.2	221	14	W83952
C	67	16	3.2	226	9	AV416808
	68	16	3.2	232	12	BI042670
	69	16	3.2	238	14	CA958302
	70	16	3.2	239	13	EX257343
	71	16	3.2	243	10	BB148038
C	72	15	3.0	60	29	AL754758
	73	15	3.0	107	12	BI683271
	74	15	3.0	116	9	AW361611
	75	15	3.0	116	13	BU998747
C	76	15	3.0	121	14	CB929267
C	77	15	3.0	123	12	BJ211548
				125	28	BH192273

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 14:45:07 ; Search time 233 Seconds
(without alignments)
5804.371 Million cell updates/sec

Title: US-09-484-577A-3

Perfect score: 501

Sequence: 1 actctccagcctctaccga.....cagaatgcgcatgatcat 501

Scoring table: OLIGO_NUC

Gapop_60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3107778

Minimum DB seq length: 0

Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Database : N_Geneseq_19Jun03.*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA.982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA.983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	4.8	60	AA0321135	Sau3AI DNA fragmen
2	24	4.8	60	AA0321135	Sau3AI DNA fragmen
3	24	4.8	60	AA0321137	Sau3AI DNA fragmen
4	24	4.8	60	AA0321137	Sau3AI DNA fragmen
5	21	4.2	24	AA065664	Primer/adaptor use
6	21	4.2	24	AA065664	Primer R Bam 24 us
7	21	4.2	24	AA0303563	Probe RBam24 for s
8	21	4.2	24	AA024707	Adaptor sequence r

Primer R-Bam-24 fo
Oligonucleotide #7
Oligonucleotide SE
Oligonucleotide SE
RDA primer R Bam24
Primer R Bam24, se
Sau3AI DNA fragmen
Nucleotide sequenc
Adaptor primer RA2
Adaptor primer RC2
Adaptor primer RA2
Adaptor primer RC2
Probe RMsp24 for s
Primer used in pro
Primer used in pro
Primer used in pro
Primer used in pro
Primer RA24 used i
Primer RA24-U used
Primer RC24-U used
Adaptor sequence r
Human cIR BspHI-Ec
Yeast PCR primer #
Yeast PCR primer #
Yeast PCR primer #
Yeast PCR primer #
Linker RA23. Sacc
Yeast PCR primer R
Yeast PCR primer R
Yeast PCR primer R
Sequencing and PCR
Sequencing and PCR
Tsp 509 I adapter
Hind III adapter p
Adapter primer nuc
Adapter primer nuc
Adapter primer nuc
Adapter primer nuc
Oligonucleotide RX
Sample adapter pri
Sample adapter pri
Rtaqi adapter, str
Sau3A I enzyme amp
BsaW I enzyme ampl
Tsp 509 I restrict
Hind III restricti
RA24 PCR primer us
RC24 PCR primer us
RC24-U PCR primer
RC24-U PCR primer
Antisense RNA prob
Antisense RNA prob
Non-biotinylated p
Primer/adaptor use
HGBV DNA PCR prime
Primer, Lu3, used
Primer R Bgl 24 us
Human Fanconi anae
Representational d
PCR suppression pr
Adaptor sequence r
Adaptor sequence r
Oligonucleotide RB
Representational d
Primer R-Bgl-24 fo
Oligo for subtrac
R Bgl 24 PCR prime
RDA Oligonucleotid
Murine H3 gene pr
Oligonucleotide pr
PCR primer R-Bgl-2

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2003, 15:30:02 ; Search time 64 Seconds
(without alignments)
3455.202 Million cell updates/sec

Title: US-09-484-577A-3
Perfect score: 501
Sequence: 1 actctcagcctctaccga.....cagaatgcgcatgatcat 501

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 874954

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCRUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY:ES

Result No.	Score	Query Match	Length	ID	Description
1	24	4.8	60	4	US-09-631-349A-19
2	24	4.8	60	4	US-09-631-349A-19
3	24	4.8	60	4	US-09-631-349A-21
4	24	4.8	60	4	US-09-631-349A-21
5	21	4.2	24	1	US-07-974-447-7
6	21	4.2	24	1	US-08-149-199-7
7	21	4.2	24	2	US-08-858-767-1
8	21	4.2	24	2	US-08-858-767-3
9	21	4.2	24	2	US-08-863-028-1
10	21	4.2	24	2	US-08-863-028-3
11	21	4.2	24	3	US-09-115-061-7
12	21	4.2	24	3	US-09-261-079-7
13	21	4.2	24	4	US-09-631-349A-1
14	21	4.2	28	2	US-08-547-214-55
15	21	4.2	28	2	US-08-663-823B-55
16	21	4.2	28	3	US-08-942-406-55
17	21	4.2	28	3	US-08-332-617-55
18	21	4.2	28	4	US-09-751-561-55
19	21	4.2	28	4	US-09-724-385-55
20	21	4.2	28	4	US-09-757-528-55
21	20	4.0	24	2	US-08-547-214-1
22	20	4.0	24	2	US-08-547-214-5
23	20	4.0	24	2	US-08-663-823B-1
24	20	4.0	24	2	US-08-663-823B-15
25	20	4.0	24	2	US-08-663-823B-60
26	20	4.0	24	2	US-08-663-823B-62
27	20	4.0	24	3	US-08-874-825-27

28	20	4.0	24	3	US-08-874-825-29	Sequence 29, Appl
29	20	4.0	24	3	US-08-874-825-31	Sequence 31, Appl
30	20	4.0	24	3	US-08-874-825-33	Sequence 33, Appl
31	20	4.0	24	3	US-08-874-825-53	Sequence 53, Appl
32	20	4.0	24	3	US-08-874-825-67	Sequence 67, Appl
33	20	4.0	24	3	US-08-874-825-107	Sequence 107, Appl
34	20	4.0	24	3	US-08-874-825-109	Sequence 109, Appl
35	20	4.0	24	3	US-08-663-824-27	Sequence 27, Appl
36	20	4.0	24	3	US-08-663-824-29	Sequence 29, Appl
37	20	4.0	24	3	US-08-663-824-31	Sequence 31, Appl
38	20	4.0	24	3	US-08-663-824-33	Sequence 33, Appl
39	20	4.0	24	3	US-08-663-824-53	Sequence 53, Appl
40	20	4.0	24	3	US-08-663-824-67	Sequence 67, Appl
41	20	4.0	24	3	US-08-663-824-107	Sequence 107, Appl
42	20	4.0	24	3	US-08-663-824-109	Sequence 109, Appl
43	20	4.0	24	3	US-08-942-406-1	Sequence 1, Appl
44	20	4.0	24	3	US-08-942-406-15	Sequence 15, Appl
45	20	4.0	24	3	US-09-381-779-3	Sequence 3, Appl
46	20	4.0	24	3	US-09-322-617-1	Sequence 1, Appl
47	20	4.0	24	3	US-09-322-617-15	Sequence 15, Appl
48	20	4.0	24	4	US-09-203-211B-1	Sequence 1, Appl
49	20	4.0	24	4	US-09-203-211B-5	Sequence 5, Appl
50	20	4.0	24	4	US-09-203-231B-19	Sequence 19, Appl
51	20	4.0	24	4	US-09-231-303-27	Sequence 27, Appl
52	20	4.0	24	4	US-09-231-303-29	Sequence 29, Appl
53	20	4.0	24	4	US-09-231-303-31	Sequence 31, Appl
54	20	4.0	24	4	US-09-231-303-33	Sequence 33, Appl
55	20	4.0	24	4	US-09-231-303-53	Sequence 53, Appl
56	20	4.0	24	4	US-09-231-303-67	Sequence 67, Appl
57	20	4.0	24	4	US-09-231-303-107	Sequence 107, Appl
58	20	4.0	24	4	US-09-231-303-109	Sequence 109, Appl
59	20	4.0	24	4	US-09-751-561-1	Sequence 1, Appl
60	20	4.0	24	4	US-09-751-561-15	Sequence 15, Appl
61	20	4.0	24	4	US-09-724-385-1	Sequence 1, Appl
62	20	4.0	24	4	US-09-724-385-15	Sequence 15, Appl
63	20	4.0	24	4	US-09-757-528-1	Sequence 1, Appl
64	20	4.0	24	4	US-09-757-528-15	Sequence 15, Appl
65	20	4.0	48	2	US-08-547-214-57	Sequence 57, Appl
66	20	4.0	48	2	US-08-663-823B-57	Sequence 57, Appl
67	20	4.0	48	3	US-08-942-406-57	Sequence 57, Appl
68	20	4.0	48	3	US-09-322-617-57	Sequence 57, Appl
69	20	4.0	48	4	US-09-751-561-57	Sequence 57, Appl
70	20	4.0	48	4	US-09-724-385-57	Sequence 57, Appl
71	20	4.0	48	4	US-09-757-528-57	Sequence 57, Appl
72	20	4.0	53	2	US-08-547-214-59	Sequence 59, Appl
73	20	4.0	53	2	US-08-663-823B-59	Sequence 59, Appl
74	20	4.0	53	3	US-08-942-406-59	Sequence 59, Appl
75	20	4.0	53	3	US-09-322-617-59	Sequence 59, Appl
76	20	4.0	53	4	US-09-751-561-59	Sequence 59, Appl
77	20	4.0	53	4	US-09-724-385-59	Sequence 59, Appl
78	20	4.0	53	4	US-09-757-528-59	Sequence 59, Appl
79	19	3.8	24	1	US-07-974-447-1	Sequence 1, Appl
80	19	3.8	24	1	US-08-149-199-1	Sequence 1, Appl
81	19	3.8	24	1	US-08-381-572-13	Sequence 13, Appl
82	19	3.8	24	1	US-08-592-820-13	Sequence 13, Appl
83	19	3.8	24	2	US-08-993-118-11	Sequence 11, Appl
84	19	3.8	24	3	US-08-748-130-23	Sequence 23, Appl
85	19	3.8	24	3	US-08-845-528C-11	Sequence 11, Appl
86	19	3.8	24	3	US-09-115-061-1	Sequence 1, Appl
87	19	3.8	24	3	US-08-963-815-6	Sequence 6, Appl
88	19	3.8	24	3	US-09-120-025-6	Sequence 6, Appl
89	19	3.8	24	3	US-09-178-973B-1	Sequence 1, Appl
90	19	3.8	24	3	US-09-261-079-1	Sequence 1, Appl
91	19	3.8	24	4	US-09-419-568F-1	Sequence 1, Appl
92	19	3.8	24	4	US-09-354-243B-1	Sequence 1, Appl
93	19	3.8	24	4	US-09-710-481-6	Sequence 6, Appl
94	19	3.8	24	4	US-09-656-952-7	Sequence 7, Appl
95	19	3.8	24	4	US-08-469-260A-1	Sequence 1, Appl
96	19	3.8	24	4	US-09-066-281B-11	Sequence 11, Appl
97	19	3.8	24	4	US-09-553-875-6	Sequence 6, Appl
98	19	3.8	24	4	US-08-488-446-1	Sequence 1, Appl
99	19	3.8	24	4	US-09-768-670-6	Sequence 6, Appl
100	19	3.8	24	4	US-08-467-344A-1	Sequence 1, Appl

ALIGNMENTS

```
RESULT 1
US-09-631-349A-19
; Sequence 19, Application US/09631349A
; Patent No. 6455255
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Birkenmeyer, Larry G.
; APPLICANT: Leary, Thomas P.
; APPLICANT: Muerhoff, A. Scott
; APPLICANT: Desai, Suresh M.
; APPLICANT: Mushahwar, Isa K.
; TITLE OF INVENTION: METHOD OF PERFORMING SUBTRACTIVE
; FILE REFERENCE: 6714.US.01
; CURRENT APPLICATION NUMBER: US/09/631.349A
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29)...(32)
; OTHER INFORMATION: n = a or g or c or t/u, unknown or other at
; OTHER INFORMATION: positions 29-32
US-09-631-349A-19

Query Match      4.8%; Score 24; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-631-349A-19
1 ACTCTCCAGCCTCTCACCAGGAT 24
|||||
4 ACTCTCCAGCCTCTCACCAGGAT 27
```

```
Query Match      4.8%; Score 24; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-631-349A-21
; Sequence 21, Application US/09631349A
; Patent No. 6455255
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Birkenmeyer, Larry G.
; APPLICANT: Leary, Thomas P.
; APPLICANT: Muerhoff, A. Scott
; APPLICANT: Desai, Suresh M.
; APPLICANT: Mushahwar, Isa K.
; TITLE OF INVENTION: METHOD OF PERFORMING SUBTRACTIVE
; FILE REFERENCE: 6714.US.01
; CURRENT APPLICATION NUMBER: US/09/631.349A
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29)...(32)
; OTHER INFORMATION: n = a or g or c or t/u, unknown or other at
; OTHER INFORMATION: positions 29-32
US-09-631-349A-21

Query Match      4.8%; Score 24; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-631-349A-21/c
; Sequence 21, Application US/09631349A
; Patent No. 6455255
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Birkenmeyer, Larry G.
; APPLICANT: Leary, Thomas P.
; APPLICANT: Muerhoff, A. Scott
; APPLICANT: Desai, Suresh M.
; APPLICANT: Mushahwar, Isa K.
; TITLE OF INVENTION: METHOD OF PERFORMING SUBTRACTIVE
; FILE REFERENCE: 6714.US.01
; CURRENT APPLICATION NUMBER: US/09/631.349A
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29)...(32)
; OTHER INFORMATION: n = a or g or c or t/u, unknown or other at
; OTHER INFORMATION: positions 29-32
US-09-631-349A-21/c
```

```
RESULT 2
US-09-631-349A-19/c
; Sequence 19, Application US/09631349A
; Patent No. 6455255
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Birkenmeyer, Larry G.
; APPLICANT: Leary, Thomas P.
; APPLICANT: Muerhoff, A. Scott
; APPLICANT: Desai, Suresh M.
; APPLICANT: Mushahwar, Isa K.
; TITLE OF INVENTION: METHOD OF PERFORMING SUBTRACTIVE
; FILE REFERENCE: 6714.US.01
; CURRENT APPLICATION NUMBER: US/09/631.349A
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29)...(32)
; OTHER INFORMATION: n = a or g or c or t/u, unknown or other at
; OTHER INFORMATION: positions 29-32
US-09-631-349A-19/c
```

```
Query Match      4.8%; Score 24; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-631-349A-21/c
; Sequence 21, Application US/09631349A
; Patent No. 6455255
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Birkenmeyer, Larry G.
; APPLICANT: Leary, Thomas P.
; APPLICANT: Muerhoff, A. Scott
; APPLICANT: Desai, Suresh M.
; APPLICANT: Mushahwar, Isa K.
; TITLE OF INVENTION: METHOD OF PERFORMING SUBTRACTIVE
; FILE REFERENCE: 6714.US.01
; CURRENT APPLICATION NUMBER: US/09/631.349A
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29)...(32)
; OTHER INFORMATION: n = a or g or c or t/u, unknown or other at
; OTHER INFORMATION: positions 29-32
US-09-631-349A-21/c
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 12:36:31 ; Search time 2864 Seconds
(without alignments)
1771.227 Million cell updates/sec

Title: US-09-484-577A-4
Perfect score: 124
Sequence: 1 LSSLSPRMSAREVVAVGCK.....KAGEILIELDPFAGGVDTAT 124

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Word size: 1
Total number of hits satisfying chosen parameters: 2012964

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09484577/runat_23102003_111504_18499/app_query_fasta_1.263
-DB=GenEmbl -Qfmt=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bite -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=250
-USER=US09484577@cgn 1.1 3508 @runat_23102003_111504_18499 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -JONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htg.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	8	6.5	231	8	AF358762	AF358762 Oryza sat
C 2	7	5.6	33	6	AR035446	AR035446 Sequence
C 3	7	5.6	42	6	AR091564	AR091564 Sequence
C 4	7	5.6	42	6	AR221814	AR221814 Sequence
C 5	7	5.6	42	6	BD134792	BD134792 Method an
C 6	7	5.6	43	6	AR091560	AR091560 Sequence
C 7	7	5.6	43	6	AR221810	AR221810 Sequence
C 8	7	5.6	43	6	BD134788	BD134788 Method an
C 9	7	5.6	45	6	AR035459	AR035459 Sequence
C 10	7	5.6	49	6	AR091563	AR091563 Sequence
C 11	7	5.6	49	6	AR091569	AR091569 Sequence
C 12	7	5.6	49	6	AR221813	AR221813 Sequence
C 13	7	5.6	49	6	AR221819	AR221819 Sequence
C 14	7	5.6	49	6	BD134791	BD134791 Method an
C 15	7	5.6	49	6	BD134797	BD134797 Method an
C 16	7	5.6	50	6	AR091559	AR091559 Sequence
C 17	7	5.6	50	6	AR221809	AR221809 Sequence
C 18	7	5.6	50	6	BD134787	BD134787 Method an
C 19	7	5.6	56	6	AR091568	AR091568 Sequence
C 20	7	5.6	56	6	AR221818	AR221818 Sequence
C 21	7	5.6	56	6	BD134796	BD134796 Method an
C 22	7	5.6	59	6	AR035460	AR035460 Sequence
C 23	7	5.6	59	6	AR035461	AR035461 Sequence
C 24	7	5.6	60	6	AR231929	AR231929 Sequence
C 25	7	5.6	60	6	AR231931	AR231931 Sequence
C 26	7	5.6	60	6	AR231931	AR231931 Sequence
C 27	7	5.6	60	6	AX464727	AX464727 Sequence
C 28	7	5.6	60	6	AX464727	AX464727 Sequence
C 29	7	5.6	60	6	AX464727	AX464727 Sequence
C 30	7	5.6	60	6	AX464729	AX464729 Sequence
C 31	7	5.6	60	6	AX464729	AX464729 Sequence
C 32	7	5.6	70	6	AR035462	AR035462 Sequence
C 33	7	5.6	70	6	AR035465	AR035465 Sequence
C 34	7	5.6	72	6	AR055681	AR055681 Sequence
C 35	7	5.6	76	6	AX240936	AX240936 Sequence
C 36	7	5.6	78	6	AR140750	AR140750 Sequence
C 37	7	5.6	78	6	I93429	I93429 Sequence 27
C 38	7	5.6	78	6	I95056	I95056 Sequence 27
C 39	7	5.6	80	14	HIVLTRCAT	M69075 Human immun
C 40	7	5.6	86	6	AR054880	AR054880 Sequence
C 41	7	5.6	86	6	AR066145	AR066145 Sequence
C 42	7	5.6	116	6	AR140738	AR140738 Sequence
C 43	7	5.6	116	6	AR140746	AR140746 Sequence
C 44	7	5.6	116	6	AR140747	AR140747 Sequence
C 45	7	5.6	116	6	I93417	I93417 Sequence 15
C 46	7	5.6	116	6	I93425	I93425 Sequence 23
C 47	7	5.6	116	6	I93426	I93426 Sequence 24
C 48	7	5.6	116	6	I95044	I95044 Sequence 15
C 49	7	5.6	116	6	I95052	I95052 Sequence 23
C 50	7	5.6	116	6	I95053	I95053 Sequence 24

REFERENCE	2	(bases 1 to 231)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
-----------	---	------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 13:28:22 ; Search time 1724 Seconds
(without alignments)
1748.118 Million cell updates/sec

Title: US-09-484-577A-4
Perfect score: 124
Sequence: 1 LSSLSPRKSAREVAVGSK.....KAGEILIELDPFAGGVAVAT 124

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4672106

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-C=/cgn2_1/USPTO_spool/US94484577/runat_23102003_111504_18512/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPEIZE=500 -MINLEN=0 -MAXLEN=250
-USER=US94484577 @CEN 1.1 2810 @runat_23102003_111504_18512 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :
1: em_estba.*
2: em_esthum.*
3: em_estcin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hrc.*
9: gb_estl.*
10: gb_est2.*
11: gb_hrc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_tod.*
26: em_gss_phg.*
27: em_gss_vri.*
28: gb_gssl.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	9	7.3	180	9	AJ477791
C 2	9	7.3	225	9	AU082059
C 3	9	7.3	249	9	AU082059
C 4	8	6.5	205	12	BM594596
C 5	8	6.5	225	28	AZ578485
C 6	8	6.5	244	12	BM876055
C 7	7	5.6	29	28	AZ853380
C 8	7	5.6	49	9	AA113042
C 9	7	5.6	64	9	AA894355
C 10	7	5.6	72	10	BF633486
C 11	7	5.6	76	13	BQ094017
C 12	7	5.6	78	29	AL938849
C 13	7	5.6	83	12	BI082225
C 14	7	5.6	89	14	R65877
C 15	7	5.6	95	13	BU652348
C 16	7	5.6	102	12	BG953925
C 17	7	5.6	104	14	CA453170
C 18	7	5.6	105	9	AA839373
C 19	7	5.6	106	9	AL503786
C 20	7	5.6	107	12	BI052864
C 21	7	5.6	110	9	AJ503237
C 22	7	5.6	112	10	BG658097
C 23	7	5.6	112	14	CA044299
C 24	7	5.6	112	14	CB504908
C 25	7	5.6	113	9	AW739688
C 26	7	5.6	113	9	AW739787
C 27	7	5.6	123	12	BM884009
C 28	7	5.6	124	10	BF414974
C 29	7	5.6	124	14	R63432
C 30	7	5.6	124	28	BH620760
C 31	7	5.6	125	28	AZ920101
C 32	7	5.6	125	28	BH849615
C 33	7	5.6	127	12	BI396580
C 34	7	5.6	128	28	AZ038470
C 35	7	5.6	130	12	BM738329
C 36	7	5.6	130	13	BY288555
C 37	7	5.6	130	29	CC335874
C 38	7	5.6	131	9	AF091519
C 39	7	5.6	131	14	CB503345
C 40	7	5.6	131	28	BH628286
C 41	7	5.6	132	12	BM040114
C 42	7	5.6	132	13	BY144970
C 43	7	5.6	132	28	BH124409
C 44	7	5.6	133	14	CA778843
C 45	7	5.6	135	13	BY129163
C 46	7	5.6	135	29	BZ784047
C 47	7	5.6	136	28	BH667508
C 48	7	5.6	138	9	AF091520
C 49	7	5.6	139	10	BE500523
C 50	7	5.6	139	12	BI033900
C 51	7	5.6	140	9	AV131901
C 52	7	5.6	142	9	AT735737
C 53	7	5.6	142	12	BI033901
C 54	7	5.6	143	10	BE571857
C 55	7	5.6	143	10	BE700540
C 56	7	5.6	143	12	BI679559
C 57	7	5.6	143	13	BY136025
C 58	7	5.6	145	9	AA013190
C 59	7	5.6	145	10	B8589065
C 60	7	5.6	145	28	AZ581119
C 61	7	5.6	148	13	B0208741
C 62	7	5.6	148	28	AZ500752

```

63      149 13 BQ835854 rf58g02.Y
64      150 9 AW407072 UI-HF-BL0
65      151 14 CD017371 NXCI_110
66      152 10 BS581435
67      153 9 AV017222 AV017222
68      154 9 AV017222 AV017222
69      155 13 BQ565096 gi30e12.Y
70      156 160 29 BZ768706 SALK_1406
71      157 13 BV369814 BV369814
72      158 13 BV369814 BV369814
73      159 2 BZ700464 QV3-NN010
74      160 2 HSM095483
75      161 13 BZ546025 BZ546025
76      162 13 BZ278133 BZ278133
77      163 28 AQ314942 AQ314942
78      164 9 AA673042 AA673042
79      165 10 BF813858 BF813858
80      166 9 AL043724 AL043724
81      167 12 BM127195 BM127195
82      168 14 CD151514 CD151514
83      169 10 BF915692 BF915692
84      170 14 T07162 T07162
85      171 28 AQ184073 AQ184073
86      172 9 AA876993 AA876993
87      173 9 AW114718 AW114718
88      174 29 BZ737358 BZ737358
89      175 10 BF841452 BF841452
90      176 9 AA876752 AA876752
91      177 9 AA886051 AA886051
92      178 14 CA989883 CA989883
93      179 28 B78965 B78965
94      180 28 B88275 B88275
95      181 9 AA876995 AA876995
96      182 9 AA586327 AA586327
97      183 28 A2463283 A2463283
98      184 10 BZ045502 BZ045502
99      185 12 B1023108 B1023108
100     186 9 AJ463996 AJ463996
100     187 9 AA658468 AA658468

```

ALIGNMENTS

```

RESULT 1
AJ477791/c
LOCUS      180 bp      mRNA      linear      EST 24-MAY-2002
DEFINITION      AJ477791 S00010 Hordeum vulgare cDNA clone S000100005G08F1, mRNA
ACCESSION      AJ477791
VERSION      AJ477791.1
KEYWORDS      EST.
SOURCE      Hordeum vulgare
ORGANISM      Hordeum vulgare
REFERENCE      1 (bases 1 to 180)
AUTHORS      Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
TITLE      Barley EST's
JOURNAL      Unpublished
COMMENT      Contact: Schulman AH
              Institute of Biotechnology
              University of Helsinki
              P.O.Box 56 (Viikinkaari 6A), University of Helsinki FIN-00014,
              Finland.

```

FEATURES

```

source
1..180
    /organism="Hordeum vulgare"
    /mol_type="mRNA"
    /db_xref="taxon:4513"
    /clone="S000100005G08F1"
    /tissue_type="Developing seed"
    /dev_stage="0..3..6..9-days after pollination"
    /clone_lib="S00010"

```

```

BASE COUNT      37 a      54 c      57 g      32 t
ORIGIN
Alignment Scores:
Pred. No.:      230      Length:      180
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    7.26%      Indels: 0
DB:             9      Gaps: 0

US-09-484-577A-4 (1-124) x AJ477791 (1-180)
Qy      22 ArgAspGluLeuAlaPheLeuProAla 30
Db      46 CGGACGAACTTCCTTCCTCCCTGCT 20

RESULT 2
AJ082059/c
LOCUS      AU082059      225 bp      mRNA      linear      EST 02-APR-2002
DEFINITION      AU082059 Rice panicle at flowering stage Oryza sativa (japonica cultivar-group) cDNA clone E0309, mRNA sequence.
ACCESSION      AU082059
VERSION      AU082059.1
KEYWORDS      EST.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM      Oryza sativa (japonica cultivar-group)
REFERENCE      1 (bases 1 to 225)
AUTHORS      Sasaki,T. and Yamamoto,K.
TITLE      Rice cDNA from panicle at flowering stage
JOURNAL      Unpublished
COMMENT      Contact: Takuji Sasaki
              National Institute of Agrobiological Resources
              Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
              305-8602, Japan
              Tel: 81-298-38-7441
              Fax: 81-298-38-7466
              Email: tsasaki@affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
              PROJECT = 'RGP'.

```

FEATURES

```

source
1..225
    /organism="Oryza sativa (japonica cultivar-group)"
    /mol_type="mRNA"
    /cultivar="Nipponbare"
    /db_xref="taxon:39947"
    /clone="E0309"
    /dev_stage="flowering stage"
    /clone_lib="Rice panicle at flowering stage"
    /note="Organ: panicle; Rice cDNA from panicle at flowering stage"

BASE COUNT      28 a      89 c      67 g      41 t
ORIGIN
Alignment Scores:
Pred. No.:      281      Length:      225
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    7.26%      Indels: 0
DB:             9      Gaps: 0

US-09-484-577A-4 (1-124) x AU082059 (1-225)
Qy      85 ValGlnProLeuGluValGlyValVal 93
Db      71 GTGACGACCTCGAAGTAGGGGTGTC 45

RESULT 3
AI597979/c
LOCUS      AI597979      249 bp      mRNA      linear      EST 14-MAY-1999

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 12:35:16 : Search time 235 Seconds
(without alignments)
1424.384 Million cell updates/sec

File: US-09-484-577A-4
Perfect score: 124
Sequence: 1 LSSLSPRMKSAREYVAVGK.....KAGEILIELDPFAGGVDVAT 124

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2552756 seqs, 1349719017 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3106038

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Command line parameters:
MODEL=frame+ p2n.model -DEV=xlh
Q=/cgn2.1/USPTO.spool/US09484577/runat.23102003.111503.18487/app_query.fasta.1.263
DB=N_Geneseq_19Jun03 -QMT=fastcap -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPCL=0
LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
LIST=100 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL
OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=250
USER=US09484577.qcgn.1.1.0 &runat.23102003.111503.18487 -NCPU=6 -ICPU=3
NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREAS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	7	5.6	33	17	AAT30598	Target binding reg
2	7	5.6	36	14	AAQ36247	HIV-1 LTR region c
3	7	5.6	38	13	AAQ33271	HIV-1 LTR - corres
4	7	5.6	43	12	AAQ12253	HIV-1 LTR mutation
5	7	5.6	43	22	AAQ12253	ECORI-AlwNI linker
6	7	5.6	45	17	AAT30611	Target binding reg
7	7	5.6	47	24	AAD29073	bx11 gene 5' end c
8	7	5.6	48	12	AAQ12252	HIV-1 LTR mutation
9	7	5.6	48	12	AAQ14775	Rennin gene probe
10	7	5.6	49	22	AAC86358	Linker used to mak
11	7	5.6	50	22	AAC86351	ECORI-AlwNI linker
12	7	5.6	56	22	AAC86357	Linker used to mak
13	7	5.6	59	17	AAT30612	Target binding reg
14	7	5.6	59	17	AAT30613	Target binding reg
15	7	5.6	60	21	AAI5068	Oligonucleotide us
16	7	5.6	60	21	AAD00174	Oligo BET-039 to g
17	7	5.6	60	24	AAD32135	Sau3AI DNA fragmen
18	7	5.6	60	24	AAD32135	Sau3AI DNA fragmen
19	7	5.6	60	24	AAD32137	Sau3AI DNA fragmen
20	7	5.6	60	24	AAD32137	Sau3AI DNA fragmen
21	7	5.6	70	17	AAT30614	Target binding reg
22	7	5.6	70	17	AAT30617	Target binding reg
23	7	5.6	72	12	AAQ12254	HIV-1 LTR Spi-bind
24	7	5.6	72	17	AAT06609	HIV wild type pLI
25	7	5.6	72	20	AAV34381	Wild type pLIIC LT
26	7	5.6	76	22	AAS23493	C. albicans essent
27	7	5.6	78	18	AAT65169	Transforming growt
28	7	5.6	86	18	AAT78708	Transforming growt
29	7	5.6	104	22	ABA50342	Class 2 SELEX gene
30	7	5.6	104	22	ABA68293	Human breast cell
31	7	5.6	104	22	ABA35295	Human foetal liver
32	7	5.6	104	22	AAK15665	Probe #13761 for g
33	7	5.6	104	22	AAI23184	Human brain expres
34	7	5.6	104	22	AAI08837	Probe #13117 for g
35	7	5.6	104	22	AAI08837	Probe #8828 used t
36	7	5.6	104	23	ABS42038	Human liver single
37	7	5.6	104	24	ABS16483	Human genome-deriv
38	7	5.6	116	18	AAT65157	Transforming growt
39	7	5.6	116	18	AAT65165	Transforming growt
40	7	5.6	116	18	AAT65166	Transforming growt
41	7	5.6	125	18	AAT65156	Transforming growt
42	7	5.6	129	22	AAS30475	DNA encoding novel
43	7	5.6	129	22	AAS30476	DNA encoding novel
44	7	5.6	129	22	AAI06255	Human reproductive
45	7	5.6	129	22	AAI06256	Human reproductive
46	7	5.6	129	22	AAK90424	Human digestive sy
47	7	5.6	129	22	AAK90425	Human digestive sy
48	7	5.6	129	22	AAI27999	Probe #17932 for g
49	7	5.6	129	24	ABN97468	Human NEDD-1 exon
50	7	5.6	153	20	AAV81498	Nucleic acid compr
51	7	5.6	159	21	AAC04535	Human secreted pro
52	7	5.6	167	24	ABL72791	Corn tassal-derive
53	7	5.6	168	12	AAQ13326	OrnatinE gene. pl
54	7	5.6	169	21	AAF14750	Aspergillus oryzae
55	7	5.6	174	24	ABA92779	HIV-1 promoter seq
56	7	5.6	178	22	ABA07490	Human ovarian and
57	7	5.6	178	22	AAI00412	Human reproductive
58	7	5.6	189	17	AAT07026	Immunogen DNA from
59	7	5.6	189	18	AAT67569	Coding sequence fo
60	7	5.6	190	25	ABX24072	Human GDP-mannose
61	7	5.6	200	19	AAV07529	Human LTR partial
62	7	5.6	210	24	ABQ75264	Human lung specifi
63	7	5.6	210	25	ABX27484	Humar: GDP-mannose
64	7	5.6	216	12	AAQ13832	Part of grg-1/chym
65	7	5.6	218	18	AAT91145	Strawberry fruit r
66	7	5.6	220	24	ABL60603	HIV long terminal

67 7 5.6 228 25 ABX26565 Human GDP-mannose
 68 7 5.6 235 24 ABN22177 Human OREF polynuc
 69 7 5.6 235 25 ABX49799 Bovine EST associa
 70 7 5.6 240 21 AA289627 HIV-1 3'-end DNA f
 71 7 5.6 242 15 AAQ72878 HIV-1 proviral pro
 72 7 5.6 242 18 AAQ72878 HIV LTR sequence.
 73 7 5.6 242 20 AAX17627 HIV proviral promo
 74 7 5.6 242 24 ABK83118 DNA binding molecu
 75 7 5.6 245 20 AA211438 HIV derived synthe
 76 7 5.6 246 16 AA119601 Human gene signatu
 77 7 5.6 249 21 AA11999 Murine PGK HRE der
 78 6 4.8 19 24 AAD40414 Bovine DGAT1 gene
 79 6 4.8 20 21 AA272729 Human biallelic ma
 80 6 4.8 20 21 AA66536 Dog genomic marker
 81 6 4.8 20 21 AA66598 Dog genomic marker
 82 6 4.8 21 18 AAT51136 Chromosomal locati
 83 6 4.8 21 24 AB597556 Human epoxide hydr
 84 6 4.8 21 24 AB598398 Human multidrug re
 85 6 4.8 22 24 ABK92070 Nucleotide sequenc
 86 6 4.8 22 24 ABK92070 Novel secreted pro
 87 6 4.8 22 25 AB258846 A'B zinc finger pr
 88 6 4.8 23 17 AAT10334 Anti-P aeruginosa
 89 6 4.8 23 22 AA233398 Nucleotide sequenc
 90 6 4.8 23 25 AB400925 Non-biotinylated p
 91 6 4.8 24 15 AAQ65638 Primer/adaptor use
 92 6 4.8 24 15 AAQ65644 Primer/adaptor use
 93 6 4.8 24 15 AAQ78833 HCMV Ab lambda lig
 94 6 4.8 24 16 AAT00012 HGBV DNA PCR prime
 95 6 4.8 24 17 AAT39259 Primer, J43, used
 96 6 4.8 24 17 AAT13195 Primer R Bgl 24 us
 97 6 4.8 24 17 AAT13201 Primer R Bam 24 us
 98 6 4.8 24 18 AAV06189 Primer used when o
 99 6 4.8 24 18 AAT92397 Adapter primer RA2
 100 6 4.8 24 18 AAT92398 Adapter primer RC2

ALIGNMENTS

RESULT 1
 ID AAT30598 standard; DNA; 33 BP.
 AC AAT30598;
 XX
 XX 19-FEB-1997 (first entry)
 DT
 DT Target binding region #18.
 DE
 DE Target binding assembly; nucleic acid recognition unit; NF-kappa-B; SPI;
 KW TATA; human papillomavirus; HPV; HPV E2; human immunodeficiency virus;
 KW HIV LTR; Tat binding unit; probe; assembly sequence; asymmetry sequence;
 KW nuclear localisation signal sequence; human cell; HPV LTR; HIV; bacteria;
 virus; ss.
 KW
 XX Synthetic.
 OS
 OS WO9617956-A2.
 PN
 PN 13-JUN-1996.
 PD
 PD 07-DEC-1995; 95WO-US15944.
 PF
 PF 09-DEC-1994; 94US-0353476.
 PR
 PR (GENE-) GENE POOL INC.
 PA Weininger AM, Weininger S;
 PA
 PI WPI; 1996-287195/29.
 XX
 XX Probe nucleic acids, target binding assemblies, etc - for detection
 XX of localisation of specific nucleic acid sequences, esp. HIV and
 PT HPV

XX Disclosure; Page 69; 172pp; English.
 XX
 CC AAT30581-730614 represent target binding regions (TBR) of a probe of the
 CC invention. The probe of the invention contains a TBR, a booster binding
 CC region (BBR), and an optional support or attachment (OSA). The target
 CC binding assembly (TBA) recognised by the probe, contains at least one
 CC nucleic acid recognition unit (NAR) and optionally a linker sequence.
 CC an assembly sequence (see AAR95994-R95998), an asymmetry sequence (see
 CC AAR95999-R96006), a nuclear localisation signal sequence (see AAR96007),
 CC and an OSA. The assembly sequence and asymmetry sequences are
 CC responsible for the folding and association of the NARs. The NARs (see
 CC AAR95965-R95993) are selected from NF-kappa-B, SPI, TATA, human
 CC papillomavirus (HPV) E2, HPV LTR, human immunodeficiency virus (HIV) LTR
 CC and Tat binding units. The linker sequence is an oligopeptide, which
 CC does not interfere with NAR function, but provides stability and control
 CC over the spacing of the NAR from the rest of the TBA. The OSA is an
 CC attached support or indicator, or other means of localisation of the
 CC probe. The probe can be used in a method for detecting or localising a
 CC specific target nucleic acid sequence (TNA). The method is highly
 CC sensitive, and has a high degree of specificity. The method can be used
 CC for detecting specific nucleic acid sequences, including those found in
 CC human cells, in HIV, HPV, and other nucleic acid containing systems,
 CC including bacteria and viruses.
 XX
 SQ Sequence 33 BP; 2 A; 9 C; 17 G; 5 T; 0 other;
 Alignment Scores: 102 Length: 33
 Pred. No.: 7.00 Matches: 7
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 5.65% Gaps: 0
 DB: 17
 US-09-484-577A-4 (1-124) x AAT30598 (1-33)
 QY 58 ValAlaTrpAlaGlyLeuGly 64
 DB 2 GTGGCTGGGGGGGACTGGGG 22
 RESULT 2
 AAQ36247
 ID AAQ36247 standard; DNA; 36 BP.
 XX
 AC AAQ36247;
 XX
 XX 25-MAR-2003 (updated)
 DT 07-JUN-1993 (first entry)
 DT
 DE HIV-1 LTR region comprising binding site for Spl-line transcription
 DE activator.
 DE
 KW Human immunodeficiency virus; AIDS; transcription initiation;
 KW long terminal repeat; target; duplex; ss.
 XX
 OS Synthetic.
 XX
 PN US5176996-A.
 PN
 PD 05-JAN-1993.
 PD
 PF 22-DEC-1989; 89US-0453532.
 PF
 PR 20-DEC-1988; 88US-0287359.
 PR 22-DEC-1989; 89US-0453532.
 XX
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 PA
 PA Hogan ME, Kessler DJ;
 PI
 PI WPI; 1993-035718/04.
 DR
 DR
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 14:16:12 ; Search time 236 Seconds

(without alignments)
1409.082 Million cell updates/sec

Title: US-09-484-577A-4

Perfect score: 124

Sequence: 1 LSSLSPRMSAREVVAVGK.....KAGEILIELDFAGGVAVAT 124

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1792395 seqs, 1340900451 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1463635

Minimum DB seq length: 0

Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Command line parameters:
-MODE=frame+p2n.model -DEV=xlh
-DB=/cgn2_1/USPRO_spo3/US09484577/runat_23102003_111506_18614/app_query.fasta.1.263
-DB=Published Applications NA -QMT=fastap -SUFFIX=oligo.rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -List=100 -DOCALIGN=200 -THR SCORE=quality -THR_MIN=1
-ALIGN=75 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=250 -USER=US09484577 @CGN 1.1 285 @runat 23102003_111506_18614 -NCPU=6
-ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRAADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTJS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	%				
No.	Score	Match Length DB ID	Description			
1	7	5.6	33	8	US-08-860-844-18	Sequence 18, Appl

33	12	US-10-407-543-18	Sequence 18, Appl
42	13	US-10-192-085-17	Sequence 17, Appl
43	13	US-10-192-085-13	Sequence 13, Appl
45	8	US-08-860-844-31	Sequence 31, Appl
45	12	US-10-407-543-31	Sequence 31, Appl
45	13	US-10-192-085-16	Sequence 16, Appl
49	13	US-10-192-085-22	Sequence 22, Appl
50	13	US-10-192-085-12	Sequence 12, Appl
56	13	US-10-192-085-21	Sequence 21, Appl
59	8	US-08-860-844-32	Sequence 32, Appl
59	8	US-08-860-844-33	Sequence 33, Appl
59	12	US-10-407-543-32	Sequence 32, Appl
59	12	US-10-407-543-33	Sequence 33, Appl
60	10	US-09-832-659-5	Sequence 5, Appl
60	11	US-09-832-658-8	Sequence 8, Appl
70	8	US-08-860-844-34	Sequence 34, Appl
70	8	US-08-860-844-37	Sequence 37, Appl
70	12	US-10-407-543-34	Sequence 34, Appl
70	12	US-10-407-543-37	Sequence 37, Appl
78	12	US-10-223-666-27	Sequence 27, Appl
104	9	US-09-864-761-20615	Sequence 20615, A
108	12	US-10-029-386-18332	Sequence 18332, A
116	12	US-10-223-666-15	Sequence 15, Appl
116	12	US-10-223-666-23	Sequence 23, Appl
116	12	US-10-223-666-24	Sequence 24, Appl
125	12	US-10-223-666-14	Sequence 14, Appl
129	9	US-09-864-761-33275	Sequence 33275, A
129	10	US-09-872-462-15	Sequence 15, Appl
129	11	US-09-764-891-8943	Sequence 8943, Ap
129	11	US-09-764-891-8944	Sequence 8944, Ap
138	12	US-10-029-386-21849	Sequence 21849, A
167	9	US-09-294-093B-2165	Sequence 2165, Ap
178	11	US-09-764-891-413	Sequence 413, App
178	14	US-10-205-428-47	Sequence 47, Appl
190	10	US-09-878-574-6131	Sequence 6131, Ap
210	10	US-09-878-574-9543	Sequence 9543, Ap
210	13	US-10-001-857-3	Sequence 3, Appl
228	10	US-09-878-574-8624	Sequence 8624, Ap
235	10	US-09-960-352-14964	Sequence 14964, A
242	11	US-09-993-346-627	Sequence 627, App
19	12	US-10-251-117-809	Sequence 809, App
19	12	US-10-251-117-1116	Sequence 1116, App
22	9	US-09-772-105-37	Sequence 37, Appl
23	14	US-09-755-398A-2	Sequence 2, Appl
23	14	US-10-191-438-2	Sequence 2, Appl
23	14	US-10-191-438-4	Sequence 4, Appl
24	8	US-08-424-550B-1	Sequence 1, Appl
24	9	US-09-751-561-1	Sequence 1, Appl
24	9	US-09-751-561-15	Sequence 15, Appl
24	9	US-09-751-797-1	Sequence 1, Appl
24	10	US-09-928-457-55	Sequence 55, Appl
24	10	US-09-928-457-58	Sequence 58, Appl
24	11	US-09-989-364-1	Sequence 1, Appl
24	11	US-09-989-364-5	Sequence 5, Appl
24	11	US-09-989-364-19	Sequence 19, Appl
24	12	US-10-096-534-69	Sequence 69, Appl
24	12	US-10-134-345-7	Sequence 7, Appl
24	12	US-10-348-190-22	Sequence 22, Appl
24	12	US-10-160-237-11	Sequence 11, Appl
24	13	US-10-032-626-7	Sequence 7, Appl
24	13	US-10-067-813-6	Sequence 6, Appl
24	13	US-10-085-108-11	Sequence 11, Appl
24	14	US-10-235-264-4	Sequence 4, Appl
24	14	US-10-193-451A-7	Sequence 7, Appl
24	14	US-10-193-451A-13	Sequence 13, Appl
24	14	US-10-193-451A-19	Sequence 19, Appl
24	14	US-10-193-451A-25	Sequence 25, Appl
24	14	US-10-001-670-27	Sequence 27, Appl
24	14	US-10-001-670-29	Sequence 29, Appl
24	14	US-10-001-670-31	Sequence 31, Appl
24	14	US-10-001-670-33	Sequence 33, Appl
24	14	US-10-001-670-53	Sequence 53, Appl

Sequence 67, Appl
Sequence 107, App
Sequence 109, App
Sequence 7993, Ap
Sequence 7919, Ap
Sequence 31415, A
Sequence 35846, A
Sequence 36091, A
Sequence 45424, A
Sequence 59601, A
Sequence 59602, A
Sequence 85954, A
Sequence 86261, A
Sequence 97035, A
Sequence 97036, A
Sequence 109737, A
Sequence 130586, A
Sequence 2, Appli
Sequence 3, Appli
Sequence 6, Appli
Sequence 7, Appli
Sequence 8, Appli
Sequence 9, Appli
Sequence 6, Appli
Sequence 97, Appli

ALIGNMENTS

RESULT 1
US-08-860-844-18
Sequence 18, Application US/08860844
Publication No. US2003C104361A1
GENERAL INFORMATION:
APPLICANT: Weininger, Susan
APPLICANT: Weininger, Arthur M
TITLE OF INVENTION: METHOD OF DETECTION OF NUCLEIC ACIDS WITH A
TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION
NUMBER OF SEQUENCES: 118
CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st St., Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/860,844
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/353,476
FILING DATE: 09-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: GP-100C1
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear

Sequence 67, Appl
Sequence 107, App
Sequence 109, App
Sequence 7993, Ap
Sequence 7919, Ap
Sequence 31415, A
Sequence 35846, A
Sequence 36091, A
Sequence 45424, A
Sequence 59601, A
Sequence 59602, A
Sequence 85954, A
Sequence 86261, A
Sequence 97035, A
Sequence 97036, A
Sequence 109737, A
Sequence 130586, A
Sequence 2, Appli
Sequence 3, Appli
Sequence 6, Appli
Sequence 7, Appli
Sequence 8, Appli
Sequence 9, Appli
Sequence 6, Appli
Sequence 97, Appli

MOLECULE TYPE: cdNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-860-844-18
Alignment Scores:
Pred No.: 47
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.65%
DB: 8
Length: 33
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0
US-09-484-577A-4 (1-124) x US-08-860-844-18 (1-33)
QY 58 ValalaTrpAlaGlyLeuGly 64
DB 2 GTGGCTGGGCGGACTGGG 22
RESULT 2
US-10-407-543-18
Sequence 18, Application US/10407543
Publication No. US20030175789A1
GENERAL INFORMATION:
APPLICANT: Weininger, Susan
APPLICANT: Weininger, Arthur M
TITLE OF INVENTION: METHOD OF DETECTION OF NUCLEIC ACIDS WITH A
TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION
NUMBER OF SEQUENCES: 118
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st St., Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/407,543
FILING DATE: 03-Apr-2003
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/860,844
FILING DATE: 09-JUN-1997
APPLICATION NUMBER: 08/353,476
FILING DATE: 09-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: GP-100C1
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cdNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-407-543-18
Alignment Scores:
Pred. No.: 47
Score: 7.00
Percent Similarity: 100.00%
Length: 33
Matches: 7
Conservative: 0

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.
OM nucleic - nuc:elic search, using sw model
Run on: October 23, 2003, 16:14:18 : Search time 223 Seconds
(without alignments)
6025.033 Million cell updates/sec

Title: US-09-484-577A-3
Perfect score: 501
Sequence: 1 actctccagccttcacga.....cagaatggcgatgatcat 501

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1792395 seqs, 1340900451 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1469544

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Database : Published Applications, NA:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:
16: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:
17: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	21	4.2	24	10	US-09-928-457-55
2	21	4.2	24	14	US-10-193-451A-13
3	21	4.2	28	9	US-09-751-561-55
4	20	4.0	23	9	US-09-755-398A-2
5	20	4.0	24	9	US-09-751-561-1
6	20	4.0	24	9	US-09-751-561-15
7	20	4.0	24	10	US-09-928-457-58
8	20	4.0	24	11	US-09-989-364-1
9	20	4.0	24	11	US-09-989-364-5
10	20	4.0	24	11	US-09-989-364-19
11	20	4.0	24	14	US-10-193-45A-25
12	20	4.0	24	14	US-10-001-670-27
13	20	4.0	24	14	US-10-001-670-29
14	20	4.0	24	14	US-10-001-670-31
15	20	4.0	24	14	US-10-001-670-33
16	20	4.0	24	14	US-10-001-670-53

24	14	US-10-001-670-67	Sequence 67, Appl
24	14	US-10-001-670-107	Sequence 107, Appl
24	14	US-10-001-670-109	Sequence 109, Appl
48	9	US-09-751-561-57	Sequence 57, Appl
53	9	US-09-751-561-59	Sequence 59, Appl
23	14	US-10-191-438-2	Sequence 2, Appli
23	14	US-10-191-438-4	Sequence 4, Appli
24	8	US-08-424-550B-1	Sequence 1, Appli
24	9	US-09-751-797-1	Sequence 1, Appli
24	11	US-09-930-334-7	Sequence 7, Appli
24	12	US-10-096-534-69	Sequence 69, Appl
24	12	US-10-134-345-7	Sequence 7, Appli
24	12	US-10-348-190-22	Sequence 22, Appl
24	12	US-10-160-237-11	Sequence 11, Appl
24	13	US-10-032-626-7	Sequence 7, Appli
24	13	US-10-067-813-6	Sequence 6, Appli
24	13	US-10-085-108-11	Sequence 11, Appl
24	14	US-10-235-264-4	Sequence 4, Appli
24	14	US-10-193-451A-7	Sequence 7, Appli
24	14	US-10-193-451A-19	Sequence 19, Appl
60	12	US-09-908-975-5483	Sequence 5483, Ap
60	12	US-09-908-975-18740	Sequence 18740, A
124	14	US-10-191-438-3	Sequence 3, Appli
180	9	US-09-791-244-3	Sequence 8, Appli
20	9	US-09-758-735-8	Sequence 12, Appl
20	9	US-09-758-735-12	Sequence 12, Appl
60	12	US-09-908-975-18749	Sequence 18749, A
103	10	US-09-954-456-300	Sequence 300, App
103	10	US-09-954-456-1201	Sequence 1201, Ap
103	10	US-09-880-107-486	Sequence 486, App
103	10	US-09-954-531-1007	Sequence 1007, Ap
103	12	US-09-873-319-94	Sequence 94, Appl
103	12	US-09-960-706-145	Sequence 145, Appl
17	12	US-10-230-006-57	Sequence 57, Appl
17	12	US-10-230-006-2103	Sequence 2103, Ap
17	12	US-10-230-006-568	Sequence 568, App
17	12	US-10-230-006-1272	Sequence 1272, Ap
19	12	US-10-226-992-23	Sequence 23, Appl
19	12	US-10-226-992-106	Sequence 106, App
25	14	US-10-098-263B-39494	Sequence 39494, A
25	14	US-10-098-263B-96190	Sequence 96190, A
87	9	US-09-864-761-29443	Sequence 29443, A
179	9	US-09-864-761-27258	Sequence 27258, A
237	10	US-09-728-444-257	Sequence 257, App
237	10	US-09-796-692-3378	Sequence 3378, App
237	10	US-10-040-862-3378	Sequence 3378, Ap
237	14	US-10-040-862-3378	Sequence 3378, Ap
248	10	US-09-983-965-3631	Sequence 3631, Ap
249	10	US-09-878-574-11290	Sequence 11290, A
15	9	US-09-755-398A-4	Sequence 4, Appli
25	14	US-10-098-263B-75929	Sequence 75929, A
25	14	US-10-098-263B-106023	Sequence 106023, A
29	12	US-10-301-516-6	Sequence 6, Appli
33	10	US-09-861-097-8	Sequence 8, Appli
33	11	US-09-861-098-8	Sequence 8, Appli
33	12	US-09-861-012-8	Sequence 8, Appli
33	13	US-10-051-989-8	Sequence 8, Appli
34	11	US-09-988-115A-67	Sequence 67, Appl
38	12	US-10-280-137-141	Sequence 141, App
38	12	US-10-280-137-142	Sequence 142, App
38	12	US-10-280-137-147	Sequence 147, App
38	12	US-10-280-137-148	Sequence 148, App
40	12	US-10-202-896-47	Sequence 47, Appl
40	12	US-10-202-896-53	Sequence 53, Appl
60	10	US-09-975-132A-8	Sequence 8, Appli
85	9	US-09-864-761-20883	Sequence 20883, A
88	10	US-09-796-692-5003	Sequence 5003, Ap
88	14	US-10-040-862-5003	Sequence 5003, Ap
105	10	US-09-998-598-1860	Sequence 1860, Ap
119	9	US-09-864-761-22266	Sequence 22266, A
137	9	US-09-864-761-21451	Sequence 21451, A
139	9	US-09-864-761-21357	Sequence 21357, A
143	9	US-09-923-876-1171	Sequence 1171, Ap
143	9	US-09-864-761-23978	Sequence 23978, A

C 90 14 2.8 156 9 US-09-736-969A-65 Sequence 65, Appl
C 91 14 2.8 157 10 US-09-878-574-8329 Sequence 8329, Ap
C 92 14 2.8 179 10 US-09-867-701-9404 Sequence 9404, Ap
C 93 14 2.8 180 12 US-10-029-366-19181 Sequence 19181, A
C 94 14 2.8 181 9 US-09-864-761-21030 Sequence 21030, A
C 95 14 2.8 190 10 US-09-960-352-9546 Sequence 9546, Ap
C 96 14 2.8 198 10 US-09-796-692-7865 Sequence 7865, Ap
C 97 14 2.8 198 10 US-09-796-692-8040 Sequence 8040, Ap
C 98 14 2.8 198 14 US-10-040-862-7865 Sequence 7865, Ap
C 99 14 2.8 198 14 US-10-040-862-8040 Sequence 8040, Ap
100 14 2.8 204 10 US-09-738-626-1469 Sequence 1469, Ap

ALIGNMENTS

RESULT 1
US-09-928-457-55
; Sequence 55, Application US/03928457
; Patent No. US20020164603A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA, specific proteins and peptides
; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
; TITLE OF INVENTION: for obtaining them and their biological application.
; NUMBER OF SEQUENCES: 99
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (OSB)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/928,457
; FILING DATE: 2001-08-14
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/214,759
; FILING DATE: 199-12-10
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-928-457-55

Query Match 4.2%; Score 21; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 ACTCTCCAGCCTCTCACCAG 21
|||
Db 4 ACTCTCCAGCCTCTCACCAG 24

RESULT 2
US-10-193-451A-13
; Sequence 13, Application US/10193451A
; Publication No. US20030096269A1
; GENERAL INFORMATION:
; APPLICANT: CULLIS, CHRISTOPHER A.
; APPLICANT: RADEMAN, SAMANTHA
; APPLICANT: KUNERT, KARL
; TITLE OF INVENTION: METHOD FOR FINDING GENETIC MARKERS
; FILE REFERENCE: 0281560005
; CURRENT APPLICATION NUMBER: US/10/193,451A
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/292,646
; PRIOR FILING DATE: 1999-04-15
; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: WordPerfect for Windows v. 7.0
; SEQ ID NO 13
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY:
; LOCATION:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
US-10-193-451A-13

Query Match 4.2%; Score 21; DB 14; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 ACTCTCCAGCCTCTCACCAG 21
|||
Db 4 ACTCTCCAGCCTCTCACCAG 24

RESULT 3
US-09-751-561-55
; Sequence 55, Application US/09751561
; Patent No. US20010007985A1
; GENERAL INFORMATION:
; APPLICANT: Rothberg, Jonathan
; APPLICANT: Deem, Michael
; APPLICANT: Simpson, John
; TITLE OF INVENTION: Method for the Determination and
; TITLE OF INVENTION: Classification of DNA Sequences in a Sample Without
; TITLE OF INVENTION: Sequencing
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie and Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/751,561
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/547,214
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mastrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-015-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)-790-9090
; TELEFAX: (212)-869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-09-751-561-55

Query Match 4.2%; Score 21; DB 9; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 60, Appl
Sequence 62, Appl
Sequence 11, Appl
Sequence 23, Appl
Sequence 11, Appl
Sequence 27, Appl
Sequence 29, Appl
Sequence 31, Appl
Sequence 33, Appl
Sequence 37, Appl
Sequence 107, Appl
Sequence 109, Appl
Sequence 27, Appl
Sequence 29, Appl

US-08-663-823B-60
US-08-663-823B-62
US-08-993-118-11
US-08-748-130-23
US-08-845-528C-11
US-08-874-825-27
US-08-874-825-29
US-08-874-825-31
US-08-874-825-33
US-08-874-825-53
US-08-874-825-67
US-08-874-825-107
US-08-874-825-109
US-08-663-824-27
US-08-663-824-29

ALIGNMENTS

RESULT 1
US-08-353-476-18
; Sequence 18, Application US/08353476
; Patent No. 5871902
; GENERAL INFORMATION:
; APPLICANT: Weininger, Susan
; APPLICANT: Weininger, Arthur M
; TITLE OF INVENTION: METHOD OF DETECTION OF DNA WITH A
; TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st St., Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/353.476
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: GP-100
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO

Alignment Scores:
Pred. No.: 21.4 Length: 33
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.65% Indels: 0
DB: 2 Gaps: 0

US-09-484-577A-4 (1-124) x US-08-353-476-18 (1-33)

QY 58 ValalaTrpAlaGlyLeuGly 64
DB 2 GTGGCTGGCGGACTGGG 22

RESULT 2
US-08-989-394-17/c
; Sequence 17, Application: US/08989394
; Patent No. 5994136
; GENERAL INFORMATION:
; APPLICANT: Naldini, Luigi
; APPLICANT: Dull, Thomas
; APPLICANT: Farson, Deborah A.
; APPLICANT: Witt, Rochelle
; TITLE OF INVENTION: METHOD AND MEANS FOR PRODUCING HIGH
; TITLE OF INVENTION: TITER, SAFE, RECOMBINANT LENTIVIRUS VECTORS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUCHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 12-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Dean H.
REGISTRATION NUMBER: 33,981
REFERENCE/DOCKET NUMBER: A7086
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA

US-08-989-394-17
Alignment Scores:
Pred. No.: 27 Length: 42
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.65% Indels: 0
DB: 2 Gaps: 0

US-09-484-577A-4 (1-124) x US-08-989-394-17 (1-42)

QY 58 ValalaTrpAlaGlyLeuGly 64
DB 37 GTGGCTGGCGGACTGGG 17

RESULT 3
US-09-271-365-17/c
; Sequence 17, Application US/09271365
; Patent No. 6165782
; GENERAL INFORMATION:
; APPLICANT: Naldini, Luigi
; APPLICANT: Dull, Thomas
; APPLICANT: Farson, Deborah A.
; APPLICANT: Witt, Rochelle
; TITLE OF INVENTION: METHOD AND MEANS FOR PRODUCING HIGH
; TITLE OF INVENTION: TITER, SAFE, RECOMBINANT LENTIVIRUS VECTORS